Cours de Mathématiques

ISA BTP, 1° année

13 juin 2013

Table des matières

1 Suites numériques **5**
Introduction . . . 5
1.1 Définitions . . . 5
1.1.1 Suites explicites/Suites récurrents . . . 6
1.1.2 Limite d'une suite . . . 8
1.1.3 Variations d'une suite . . . 9
1.1.4 Suites bornées . . . 10
1.1.5 Suites extraites . . . 10
1.2 Critères de convergence . . . 11
1.2.1 Les suites arithmétiques et géométriques . . . 11
1.2.2 Opérations sur les limites . . . 12
1.2.3 Comparaison . . . 13
1.2.4 Suites monotones bornées . . . 14
1.2.5 Suites extraites et convergence . . . 15
1.3 Suites récurrentes . . . 16
1.3.1 Généralités . . . 16
1.3.2 Suites récurrentes linéaires à coefficients constants . . . 16
1.3.3 Suites récurrentes d'ordre 1 . . . 18
1.4 Vitesse de convergence d'une suite récurrente . . . 22
1.4.1 Critère de convergence . . . 22
1.4.2 Échelle de vitesses . . . 23
1.4.3 Application aux suites récurrente . . . 24
1.5 Méthode de Newton . . . 25

2 Espaces vectoriels **27**
Introduction . . . 27
2.1 Le plan et l'espace . . . 28
2.1.1 Qu'est-ce qu'un vecteur ? . . . 28
2.1.2 Repères et coordonnées . . . 29
2.1.3 L'idée d'espace vectoriel . . . 30
2.2 Définitions et Exemples . . . 31
2.2.1 Loi de composition interne . . . 31
2.2.2 Multiplication externe . . . 31

2.2.3 Espace vectoriel 32
2.2.4 Sous espaces vectoriels 34
2.2.5 Sommes de sous espaces vectoriels 38
2.3 Familles de vecteurs 39
2.3.1 Familles génératrices 39
2.3.2 Familles libres 41
2.3.3 Bases et coordonnées 43
2.3.4 Dimension d'un espace vectoriel 44
2.4 Représentations matricielles 46
2.4.1 Matrices 46
2.4.2 Changement de base 48
2.5 Applications linéaires 50
2.5.1 Définition 50
2.5.2 Premières propriétés 51
2.5.3 Matrice d'une application linéaire 52
2.5.4 Image et noyau d'une application linéaire 53
2.5.5 Injectivité, surjectivité d'une application linéaire 56
2.5.6 Matrice d'une application linéaire 58
2.5.7 Diagonalisation d'une application linéaire 62

Chapitre 1

Suites numériques

Introduction

Une suite, une suite de nombres, est la donnée d'une série de nombres dans un ordre précis. En les choisissant les uns après les autres, on peut construire n'importe quelle suite de nombres. Une suite construite au hasard est une suite aléatoire. Dans ce chapitre, on va étudier des suites de nombre construites sur la logique. On va commencer par voir quelles sont les différentes façons de définir une suite de façon logique, puis on étudiera des outils permettant de déterminer le comportement d'une suite donnée ; c'est-à-dire la façon dont évoluent les différents termes de la suite quand on les parcours un à un ou encore ses variations et/ou sa limite.

1.1 Définitions

Formellement, une suite de nombres réels $(u_n)_{n\in\mathbb{N}}$ est une fonction de l'ensemble des entiers ($\mathbb{N}$) dans l'ensemble des réels ($\mathbb{R}$) :

$$\begin{array}{rccc} u & : & \mathbb{N} & \longrightarrow & \mathbb{R} \\ & & n & \longmapsto & u(n) = u_n \end{array}$$

On a l'habitude de mettre n en indice plutôt qu'entre parenthèses. On dit que $\mathbb{N}$ est l'ensemble des indices. Il peut arriver qu'une suite commence à $n = 1$ (si par exemple u_0 ne peut exister). L'ensemble des indices est alors $\mathbb{N}^*$. De façon générale, on peut faire commencer une suite de n'importe quel entier.

Si l'ensemble des indices est $\mathbb{N}$, u_0 est le premier terme de la suite, u_1 le deuxième, etc... . On notera u_n le terme général, ou terme de rang n, et l'on notera $(u_n)_{n\in\mathbb{N}}$ l'ensemble des termes de la suite.

1.1.1 Suites explicites/Suites récurrents

Il y a deux façons de construire une suite de façon logique :

1. En définissant chaque terme en fonction de sa place dans la suite ; c'est-à-dire en donnant explicitement u_n en fonction de n. C'est la forme explicite. Par exemple

   – $(u_n)_{n\in\mathbb{N}}$: $u_n = n^2 + 1$.
   – $(v_n)_{n\in\mathbb{N}^*}$: $v_n = e^{\left(\frac{1}{n}\right)}$.

   – $(w_n)_{n\in\mathbb{N}^*}$: $w_n = \sin\left(\frac{n\pi}{2}\right)$.

2. En donnant une règle logique pour passer d'un terme au suivant ; c'est à dire en donnant u_n en fonction du ou des termes précédents $(u_{n-1}, u_{n-2}, \ldots)$. C'est la forme récurrente. La relation liant u_n aux termes précédents est la relation de récurrence. Pour définir complètement une suite de façon récurrente, il faut également donner le point de départ, autrement dit le premier terme u_0 (ou u_1 si la suite commence à $n = 1$). Par exemple

$$(u_n)_{n\in\mathbb{N}} : \begin{cases} u_0 = 1 \\ u_n = u_{n-1} + 2, \qquad \forall n \geqslant 1 \end{cases}$$

$$(v_n)_{n\in\mathbb{N}} : \begin{cases} v_0 = 0, v_1 = 1 \\ v_n = v_{n-1} + v_{n-2}, \qquad \forall n \geqslant 2 \end{cases}$$

   Remarque : comme on le voit sur le second exemple, si l'on définit v_n en fonction des deux termes précédents, il faut donner les deux premières valeurs. De façon générale, il faut suffisamment de premiers termes pour pouvoir amorcer la suite.

La définition explicite est de loin la plus pratique, puisque chaque terme est accessible par un calcul direct alors que pour une suite récurrente, si l'on veut connaître un terme (par exemple le 100-ième), il faut d'abord calculer tous les termes précédents.

D'autre part, pour étudier une suite explicite définie par une relation de la forme $u_n = f(n)$, on peut utiliser les outils d'analyse que l'on connait (dérivées, limites, comparaison locale, ...).

Cependant, la forme récurrente est souvent la plus adaptée lorsque l'on utilise les suites pour représenter des situations concrètes (calculs d'intérêts, évolution d'un système mécanique, évolution d'une variable dans un programme,...).

Pour étudier de telles suites, on commence en général par en chercher une forme explicite. On verra que c'est en particulier possible dans les cas où la relation de récurrence est

linéaire.

On verra enfin comment étudier une suite récurrente dans les cas où l'on ne peut en obtenir une forme explicite.

Exemples. On connaît quelques suites récurrentes célèbres.

1. *Les suites arithmétiques.* Une suite arithmétique est une suite telle que pour passer d'un terme au suivant, on ajoute toujours la même chose. La forme récurrente d'une suite arithmétique est donc

$$(a_n) \ : \ \begin{cases} a_0 \text{ donné} \\ a_n = a_{n-1} + r. \end{cases}$$

où r est un nombre fixé (indépendant de n) que l'on appelle la raison. Les suite arithmétiques permettent par exemple de représenter l'évolution d'un capital placé à intérêts simples.

On connaît (ou on sait facilement retrouver) la forme explicite d'une suite arithmétique :

$$a_n = a_0 + n \times r$$

On sait également exprimer la somme des premiers termes d'une suite arithmétique :

$$S_n = a_0 + a_1 + \ldots + a_n = (n+1)\frac{a_0 + a_n}{2}.$$

2. *Les suites géométriques.* Une suite géométrique est une suite où pour passer d'un terme au suivant, on multiplie toujours par la même quantité q :

$$(g_n) \ : \ \begin{cases} g_0 \text{ donné} \\ g_n = q \times g_{n-1}. \end{cases}$$

On appelle encore q la raison. Les suites géométriques permettent par exemple de modéliser l'évolution d'un capital placé à intérêts composés.

La forme explicite d'une suite géométrique est

$$g_n = g_0 \times q^n.$$

D'autre part, on sait là encore exprimer la somme des premiers termes d'une suite géométrique :

$$\Sigma_n = g_0 + g_1 + \ldots + g_n = g_0 \times \frac{1 - q^{n+1}}{1 - q}.$$

3. *La suite de Fibonacci.* La suite de Fibonacci est la suite définie par

$$(F_n) : \begin{cases} F_0 = 0,\ F_1 = 1, \\ F_n = F_{n-1} + F_{n-2}. \end{cases}$$

C'est une suite récurrente d'ordre 2 (car chaque terme est donné en fonction des deux précédents). On verra comment donner une forme explicite de ce type de suite, dans certains cas particuliers. On verra en particulier que l'on peut montrer que la suite des quotients (F_{n+1}/F_n) tend vers le nombre d'or.

1.1.2 Limite d'une suite

La limite d'une suite permet de décrire le comportement d'une suite lorsque n tend vers l'infini. Par exemple, la suite (u_n) définie par

$$u_n = \frac{(-1)^n}{n}$$

a ses termes qui se rapprochent de 0 quand n grandit. On dit alors que (u_n) a pour limite 0 et on note

$$\lim_{n\to+\infty} u_n = 0.$$

Contrairement à la suite (u_n) précédente, la suite (v_n) définie par $v_n = n^2$ ne se rapproche d'aucun nombre. Les termes v_n sont de plus en plus grands. Dans ce cas, on note

$$\lim_{n\to+\infty} v_n = +\infty.$$

Enfin, si l'on observe la suite (w_n) définie par $w_n = (-1)^n$, on voit que cette suite prend alternativement les valeurs 1 et -1. Une telle suite n'a donc pas de limite.

Remarques :
- Pour une suite, on s'intéressera toujours à la limite $n \to +\infty$. On pourra alors se contenter de noter $\lim_n u_n$.
- Si une suite admet une limite finie, on dit qu'elle converge. Sinon elle diverge (même si elle tend vers $+\infty$).

Formellement, on peut alors établir les définitions suivantes :

- Une suite (u_n) admet comme limite le nombre $\ell \in \mathbb{R}$ si et seulement si

$$\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N} \text{ tq } \forall n \geqslant n_0,\ |u_n - \ell| < \varepsilon.$$

Dans cette définition, $\forall \varepsilon > 0$ se lit "pour ε aussi petit que l'on veut" et $|u_n - \ell| < \varepsilon$ se lit "la distance entre u_n et ℓ est plus petite que ε". Autrement dit, cette définition nous dit que quitte à prendre n assez grand, la suite (u_n) se rapproche de ℓ aussi près que l'on veut et reste proche de ℓ.

– Une suite (u_n) tend vers $+\infty$ si et seulement si

$$\forall A > 0,\ \exists n_0 \in \mathbb{N} \text{ tq } \forall n \geqslant n_0,\ u_n > A.$$

Exemples :

– Pour (u_n) : $u_n = \frac{1}{n}$, il est clair que plus n est grand, plus son inverse est petit. Ainsi, si l'on se fixe une quantité $\varepsilon > 0$. On pourra trouver un rang n_0 à partir duquel on a

$$\forall n \geqslant n_0, \quad \frac{1}{n} < \varepsilon.$$

C'est en particulier le cas pour n'importe quel entier $n_0 > \frac{1}{\varepsilon}$.

– Soit (u_n) une suite *d'entiers*. On peut montrer (en exploitant le fait que la distance entre deux entiers est toujours $\geqslant 1$) que si (u_n) est convergente, alors elle est stationnaire (i.e. constante à partir d'un certain rang).

– De façon générale, à partir de la définition de la convergence, on peut montrer que si une suite est convergente, alors sa limite est unique (en raisonnant par l'absurde).

1.1.3 Variations d'une suite

Une suite (u_n) sera dite

– *stationnaire* si tous ses termes sont égaux à partir d'un certain rang :

$$(u_n) \text{ est stationnaire} \iff \exists n_0 \text{ tq } \forall n \geqslant n_0,\ u_{n+1} = u_n.$$

Remarque : si pour un $n_0 \in \mathbb{N}$, on a $u_n = u_{n_0}$ pour tout $n \geqslant n_0$, on dira que (u_n) est stationnaire à partir d'un certain rang.

– *croissante* si pour tout $n \in \mathbb{N}$, on a $u_{n+1} \geqslant u_n$.
– *strictement croissante* si pour tout $n \in \mathbb{N}$, on a $u_{n+1} > u_n$.

– *décroissante* si pour tout $n \in \mathbb{N}$, on a $u_{n+1} \leqslant u_n$.
– *strictement décroissante* si pour tout $n \in \mathbb{N}$, on a $u_{n+1} < u_n$.

– *monotone* si elle est soit croissante (i.e. $\forall n \in \mathbb{N}$, $u_{n+1} \geqslant u_n$), soit décroissante (i.e. $\forall n \in \mathbb{N}$, $u_{n+1} \leqslant u_n$).

En pratique, pour déterminer les variations d'une suite, on étudie le signe de $u_{n+1} - u_n$. Si ce signe est constant (i.e. le même pour tout n), alors la suite est monotone (et le sens dépend du signe). Si ce signe dépend de n, alors la suite n'est pas monotone.

Dans le cas des suites positives, il sera parfois plus facile d'étudier le quotient $\dfrac{u_{n+1}}{u_n}$ et de le comparer à 1.

Exemples :

– La suite (u_n) : $u_n = n^3$ est croissante.
– La suite (v_n) : $v_n = \dfrac{1}{n}$ est décroissante.

Remarque : il se peut que le signe de $u_{n+1} - u_n$ dépende de n pour de petites valeurs de n, et finisse par se stabiliser. On dira alors que (u_n) est monotone *à partir d'un certain rang.*

1.1.4 Suites bornées

Une suite (u_n) sera dit
– *majorée* s'il existe un nombre M (indépendant de n) tel que

$$\forall n \in \mathbb{N},\ \ u_n \leqslant M.$$

– *minorée* s'il existe un nombre m (indépendant de n) tel que

$$\forall n \in \mathbb{N},\ \ u_n \geqslant M.$$

– *bornée* si elle est à la fois majorée et minorée ; autrement dit s'il existe $m, M \in \mathbb{R}$ tels que

$$\forall n \in \mathbb{N},\ \ m \leqslant u_n \leqslant M.$$

Pour des raisons pratiques, on utilisera plutôt la définition suivante :

$$(u_n) \text{ est bornée } \iff \exists A > 0 \text{ tq } \forall n \in \mathbb{N},\ |u_n| \leqslant A.$$

Exemples :
– La suite (u_n) : $u_n = n$ est minorée mais non majorée.
– La suite (v_n) : $v_n = -e^n$ est majorée mais non minorée.
– Les suites (w_n) : $w_n = \dfrac{1}{n}$, (x_n) : $x_n = (-1)^n$ et (y_n) : $y_n = \sin(n^2 - n + 1)$ sont bornées.
– La suite (x_n) : $x_n = n(-1)^n$ n'est ni majorée ni minorée.

1.1.5 Suites extraites

Soit (u_n) une suite réelle. Une suite extraite (ou sous-suite) de (u_n) est une suite construite à partir de certains éléments de (u_n). On peut par exemple ne prendre que les termes d'indices pairs, ou les termes dont l'indice est un nombre premier. Cependant, l'extraction ne peut pas se faire n'importe comment. Il faut en particulier respecter l'ordre donné par la suite (u_n). Une fois que l'on a extrait un terme, le suivant doit être pris parmi les termes suivants. On ne peut pas revenir en arrière. Formellement, une suite extraite est la donnée d'une fonction

$$\begin{array}{rccc} \varphi & : & \mathbb{N} & \longrightarrow & \mathbb{N} \\ & & n & \longmapsto & \varphi(n) \end{array}$$

qui doit être strictement croissante. La suite extraite correspondant est alors la suite

$$(v_n) \ : \ v_n = u_{\varphi(n)}.$$

Ainsi, la sous-suite donnée par les termes d'indices pairs correspond à la fonction $\varphi \ : \ n \mapsto 2n$.

Exemple : soit (u_n) la suite définie par $u_n = (-1)^n$. De (u_n) on peut extraire les suites d'indices pairs et d'indices impairs :

$$(v_n) \ : \ v_n = u_{2n} = 1 \quad \text{et} \quad (w_n) \ : \ w_n = u_{2n+1} = -1$$

Remarque : ce sont toutes deux des suites constantes, alors que (u_n) ne l'est pas.

La notion de suite extraite peut être un outil efficace pour étudier une suite donnée, mais il faut être très rigoureux dans son utilisation.

Ainsi, à l'image de l'exemple précédent, certaines propriétés peuvent être vérifier par des suites extraites de (u_n) mais pas par (u_n) elle même.

1.2 Critères de convergence

Comme on l'a dit plus haut, étudier une suite donnée c'est déterminer ses variations et/ou sa limite.

En pratique, il existe de nombreux outils permettant de déterminer le comportement d'une suite donnée. Dans le cas des suites explicites, on verra que l'on peut par exemple adapter les outils que l'on connait pour l'étude des fonctions.

On verra également qu'il existe de nombreux outils permettant de comparer des suites entre elles. On pourra alors établir un catalogue de suites-types dont on connaît parfaitement le comportement et auxquelles on pourra comparer les suites qui nous intéressent.

Enfin, on verra également que certains de ces outils permettent d'étudier le comportement des suites récurrentes même lorsque l'on ne peut avoir de forme explicite.

Commençons par deux exemples connus.

1.2.1 Les suites arithmétiques et géométriques

- Une suite arithmétique (a_n) de premier terme a_0 est de raison r est
  - croissante et tend vers l'infini si $r > 0$,
  - décroissante et tend vers $-\infty$ si $r < 0$,
  - constante si $r = 0$.

- Une suite géométrique (g_n) de premier terme $g_0 > 0$ et de raison q est
  - croissante et tend vers l'infini si $q > 1$,
  - constante si $q = 1$,
  - décroissante et tend vers 0 si $0 \leqslant q < 1$,
  - alternée et tend vers 0 si $-1 < q < 0$,
  - non convergente si $q \leqslant -1$.

Ces résultats s'obtiennent rapidement à partir des formes explicites des suites arithmétiques et géométriques. Or elles s'appliquent facilement à des suites récurrentes. En effet, pour vérifier qu'une suite est arithmétique, il suffit de montrer que la différence de deux termes consécutifs ne dépend pas de n.

De même, pour montrer qu'une suite donnée et géométrique, il suffit de montrer que le quotient de deux termes consécutifs est constant.

Exemple : soit (u_n) définie par

$$\forall n \in \mathbb{N}, \quad u_n = \frac{e^n}{2^{2n+1}}.$$

1. Montrer que (u_n) est une suite géométrique.
2. Donner son premier terme et sa raison.
3. En déduire les variations et la limite de (u_n).

1.2.2 Opérations sur les limites

Soient (u_n) et (v_n) deux suites convergentes et soient ℓ_1 et ℓ_2 leurs limites respectives. On peut montrer facilement les suites (s_n), (p_n) et (q_n) définies par

$$s_n = u_n + v_n, \quad p_n = u_n \times v_n, \quad q_n = \frac{u_n}{v_n}$$

sont également convergentes (avec quelques condition évidentes pour (q_n)). De plus, les limites de ces suites se déduisent facilement de ℓ_1 et ℓ_2 :

$$\begin{aligned} \forall \lambda, \mu \in \mathbb{R}, \quad \lim_n (\lambda.u_n + \mu.v_n) &= \lambda.\ell_1 + \mu.\ell_2 \\ \lim_n (u_n \times v_n) &= \ell_1 \times \ell_2 \\ \text{si } \ell_2 \neq 0, \qquad \lim_n \frac{u_n}{v_n} &= \frac{\ell_1}{\ell_2} \end{aligned}$$

Remarque : la suite (q_n) n'est définie qu'à partir du moment où $v_n \neq 0$. Or si $\ell_2 \neq 0$, on est sur qu'il existe un rang à partir duquel $v_n \neq 0$.

Enfin, de façon générale, on peut exploiter les limites de fonctions et les règles de composition.

Ces règles permettent de déterminer les limites de suites formées de somme, produits, quotients de suites connues ; il est cependant parfois nécessaire de travailler un peu sur les expressions que l'on souhaite étudier pour lever des indéterminations.

Exemples :
- Limite de $\dfrac{n^2+1}{n^2}$
- Limite de $\dfrac{n}{n^3+1}$
- Limite de $e^{\frac{1}{n}}$

1.2.3 Comparaison

Limites et inégalités

Soient (u_n) et (v_n) deux suites convergentes et soient ℓ_1 et ℓ_2 leurs limites respectives. Si $u_n \leqslant v_n$ pour tout n, alors $\ell_1 \leqslant \ell_2$.

Remarque : si $u_n < v_n$ pour tout n, on ne peut pas avoir mieux que $\ell_1 \leqslant \ell_2$. Penser à $u_n = -\dfrac{1}{n}$ et $v_n = \dfrac{1}{n}$.

De même, on a un théorème d'encadrement connu sous le nom de théorème des gendarmes :

Théorème 1.2.1 *Soient (u_n), (v_n) et (w_n) trois suites telles que*

1. $\forall n \in \mathbb{N}, \quad u_n \leqslant v_n \leqslant w_n$.
2. *(u_n) et (w_n) sont convergentes et $\lim\limits_n u_n = \lim\limits_n w_n = \ell$.*

Alors la suite (v_n) converge également et $\lim\limits_n v_n = \ell$.

Comparaison locale

Comme pour les fonctions d'une variable, il est possible de comparer les suites numériques qui sont à termes positif. Cependant, comme pour les limites, les comparaisons se font ici pour $n \to +\infty$. Ainsi, soient (u_n) et (v_n) deux suites positives.

- Elles sont dites équivalentes (en $+\infty$) si

$$\lim_{n\to+\infty} \frac{u_n}{v_n} = 1$$

et on note $u_n \underset{\infty}{\simeq} v_n$.

– On dit que (v_n) domine (u_n) si

$$\lim_{n\to+\infty} \frac{u_n}{v_n} = 0$$

et on note $u_n = o_\infty(v_n)$.

Exemples

1. Les suites (u_n) et (v_n) définies par

$$\forall n \in \mathbb{N}^*, \quad u_n = \frac{1}{n^2} \quad \text{et} \quad v_n = \frac{1}{n^2+1}$$

sont équivalentes.

2. Les suites (u_n) et (v_n) définies par

$$\forall n \in \mathbb{N}^*, \quad u_n = \frac{1}{n} \quad \text{et} \quad v_n = e^{-n}$$

vérifient $u_n = o(v_n)$.

À l'aide de ces critères de comparaison locale, on peut traiter de nombreuses suites explicites, notamment les suites définies par des fractions rationnelles.

On peut ainsi montrer que deux suites équivalentes ont le même comportement : si l'une converge, l'autre aussi et elles ont même limite.

De même, on peut montrer que si une suite (v_n) tend vers 0 alors toute suite positive dominée par (v_n) tend également vers 0.

D'autre part, ces critères de comparaison permettent de mettre en place une échelle pour mesurer la vitesse de convergence. Précisément, on peut montrer que deux suites équivalentes qui convergent tendent vers leur limite commune à la même vitesse, i.e. à une valeur de n donnée, les distances entre les termes de rang n et la limite commune ℓ sont du même ordre de grandeur.

De même, on peut montrer que si une suite (v_n) tend vers 0, alors toute suite positive dominée par (v_n) tend vers 0 plus rapidement que (v_n).

On reviendra sur tout cela en TP machines.

1.2.4 Suites monotones bornées

Il faut en général se méfier de notre intuition lorsque l'on travaille avec des suites. On a toutes fois quelques résultats qui semblent naturels, et qui le sont. Ainsi, on a le théorème suivant :

Théorème 1.2.2 *Toute suite réelle croissante et majorée est convergente. De même, toute suite décroissante et minorée est convergente.*

Remarque : ce théorème est très pratique pour déterminer la nature d'une suite. Cependant, le calcul de la limite éventuelle doit se faire avec précaution. La limite est alors la *borne optimale* de la suite étudiée...

Un autre théorème très utile concerne les *adjacentes* : on dira que deux suites (u_n) et (v_n) sont adjacentes si

1. $\forall n \in \mathbb{N}, \quad u_n \leqslant v_n$.
2. La suite (u_n) est croissante.
3. La suite (v_n) est décroissante.
4. $\lim_n (u_n - v_n) = 0$.

On a alors la propriété suivante :

Proposition 1.2.3 *Si deux suites (u_n) et (v_n) sont adjacentes, alors elles convergent et ont même limite.*

1.2.5 Suites extraites et convergence

Le lien fort qu'il existe entre une suite donnée (u_n) et les différentes suites $(u_{\varphi(n)})$ que l'on peut en extraire se traduit naturellement au niveau du comportement et de la convergence. Ainsi, on peut rapidement montrer que les propriétés de la suite (u_n) (variations, majorations, convergence,...) se transmettent aux suites extraites : si (u_n) est monotone (resp. bornée, convergente), alors toute suite extraite de (u_n) est également monotone (resp. bornée, convergente). D'autre part, si (u_n) est convergente, alors toute suite extraite de (u_n) admet la même limite que (u_n).

Cette remarque permet de développer des outils d'étude "par défaut" d'une suite donnée via quelques unes de ses suites extraites. Ainsi, si l'on peut trouver deux suites $(u_{\varphi(n)})$ et $(u_{\psi(n)})$ qui convergent mais qui n'admettent pas la même limite, alors la suite (u_n) ne peut converger.

Exemple : Soit (u_n) définie par

$$\forall n \in \mathbb{N}, \quad u_n = (-1)^n.$$

Les deux suites extraites (u_{2n}) et (u_{2n+1}) convergent respectivement vers 1 et -1. Ces limites étant différentes, on en déduit que (u_n) ne peut converger.

Intuitivement, il faut donc retenir que l'étude de quelques suite extraites d'une suite (u_n) ne permettent pas en général d'obtenir de l'information générale sur (u_n). Cependant, on peut établir le résultat suivant :

Proposition 1.2.4 *S'il on peut extraire d'une suite (u_n) deux suite (v_n) et (w_n) telles que*

1. *(v_n) et (w_n) convergent et ont même limite ℓ,*
2. *(v_n) et (w_n) recouvrent à elles deux tous les termes de (u_n),*

alors la suite (u_n) converge vers ℓ.

1.3 Suites récurrentes

1.3.1 Généralités

On appelle *suite récurrente* toute définie par

$$(u_n) \,:\, (u_n) \,:\, \begin{cases} u_0 \text{ donné} \\ u_{n+1} = f(u_n, u_{n-1}, \ldots) \end{cases}$$

Si u_{n+1} ne dépend que de u_n, on dira que (u_n) est d'ordre 1. Si u_{n+1} dépend de u_n et u_{n-1},etc... Si l'on souhaite étudier une suite récurrente (i.e. trouver ses variations, sa limite éventuelle,...) l'idéal est de trouver une forme explicite. On verra des une méthode efficace dans certains cas particuliers. En règle générale, on ne peut malheureusement pas espérer trouver de forme explicite. On a quand même tous les outils que l'on a mis en place plus haut pour étudier ces suites. D'autre part, on peut également tirer de l'information de la fonction f.

1.3.2 Suites récurrentes linéaires à coefficients constants

Suites d'ordre 1 Une suite récurrente linéaire d'ordre 1 est une suite (u_n) vérifiant

$$u_{n+1} = au_n.$$

Ce sont les suites géométriques. On connaît parfaitement le comportement d'une telle suite en fonction de la raison a.

Suites d'ordre 2 Une suite récurrente linéaire d'ordre 2 est une suite (v_n) vérifiant

$$v_{n+2} = \alpha v_{n+1} + \beta v_n, \quad \alpha, \beta \in \mathbb{R}$$

ou plus généralement

$$av_{n+2} + bv_{n+1} + cv_n = 0, \quad a, b, c \in \mathbb{R}.$$

Remarque : pour définir complètement une telle suite, il faut donner également les deux premiers termes v_0 et v_1.

Exemple : la suite de Fibonacci est par exemple une suite récurrente d'ordre 2 dont les coefficients sont $a = b = 1$ et les premiers termes sont $F_0 = 0$, $F_1 = 1$.

Il existe une méthode permettant d'obtenir la forme explicite d'une telle suite en fonction de a et b. On pourra noter que la démarche et les conclusions sont très proches de ce que l'on fait et obtient lors de la résolution d'une équation différentielle d'ordre 2 à coefficients constants. Ainsi, soit (u_n) vérifiant $au_{n+2} + bv_{n+1} + cv_n$. On appelle *l'équation caractéristique* de (v_n) l'équation

$$(E) \ : \ aX^2 + bX + c = 0$$

En notant alors $\Delta = b^2 - 4ac$ le discriminant de (E), on a

- Si $\Delta > 0$, alors (E) possède deux racines réelles α_1 et α_2 et la forme explicite de (v_n) est
$$\forall\, n \in \mathbb{N}, \quad v_n = A.\alpha_1^n + B.\alpha_2^n$$
où A et B sont des constantes fixe, déterminées par les conditions initiales (i.e. la donnée de v_0 et v_1).

- Si $\Delta = 0$, alors (E) possède une racine double α_0 et la forme explicite de (v_n) est
$$\forall\, n \in \mathbb{N}, \quad v_n = \alpha_0^n(A + n.B).$$

- Si $\Delta < 0$, (E) possède deux racines complexes $\alpha_{1,2} = re^{\pm i\theta}$. La forme explicite de (v_n) est alors de la forme
$$\forall\, n \in \mathbb{N}, \quad v_n = A.r^n \cos(n\theta) + B.r^n \sin(n\theta).$$

Dans tous les cas, pour déterminer les coefficients A et B, on utilise les conditions initiales v_0 et v_1. Si par exemple $\Delta > 0$, on a

$$v_n = A.\alpha_1^n + B.\alpha_2^n.$$

En posant $n = 0$ puis $n = 1$, on obtient un système 2×2 dont les inconnues sont A et B :

$$\begin{cases} u_0 = A + B \\ u_1 = A.\alpha_1 + B.\alpha_2 \end{cases}$$

En résolvant ce système, on trouve A et B.

Exemple : la suite de Fibonacci

$$(F_n) \ : \ \begin{cases} F_0 = 1, \ F_1 = 1 \\ F_{n+2} = F_{n+1} + F_n, \end{cases}$$

L'équation caractéristique de (F_n) est

$$X^2 - X - 1 = 0$$

dont le discriminant est

$$\Delta = 1 + 4 \times 1 = 5$$

et les solution sont

$$\text{le nombre d'or } \varphi = \frac{1+\sqrt{5}}{2} \quad \text{et} \quad \text{son inverse } \varphi^{-1} = \frac{1-\sqrt{5}}{2}.$$

La suite de Fibonacci est donc de la forme

$$F_n = A\varphi^n + B\varphi^{-n}.$$

Pour déterminer les coefficients A et B, on utilise les conditions initiales :

$$\begin{aligned} F_0 = 0 \quad &\Rightarrow \quad A + B = 0 \\ F_1 = 1 \quad &\Rightarrow \quad A.\varphi + \frac{B}{\varphi} = 1 \end{aligned}$$

D'où

$$A = -B = \frac{\varphi}{\varphi^2 - 1}.$$

On peut en particulier vérifier que si $Q_n = F_{n+1}/F_n$, la suite (Q_n) tend vers φ. (Exercice : à faire).

1.3.3 Suites récurrentes d'ordre 1

Soit (u_n) une suite récurrente d'ordre 1 :

$$(u_n) \; : \; \begin{cases} u_0 \text{ donné} \\ u_{n+1} = f(u_n) \end{cases}$$

Comme on l'a dit plus haut, l'étude de la suite (u_n) se fait via l'étude de la fonction f.

On supposera dans un premier temps que f est *croissante*.

Variations de (u_n). Voyons comment les variations de f permettent de déterminer les variations de (u_n). Puisque f est croissante, on a par définition

$$\forall a, b \qquad a \leqslant b \;\Rightarrow\; f(a) \leqslant f(b).$$

En appliquant cela à deux termes consécutifs de la suite (u_n), on a :

$$u_{n+1} \leqslant u_n \;\Rightarrow\; f(u_{n+1}) \leqslant f(u_n) \;\Rightarrow\; u_{n+2} \leqslant u_{n+1}.$$

De même,

$$u_{n+1} \geqslant u_n \;\Rightarrow\; f(u_{n+1}) \geqslant f(u_n) \;\Rightarrow\; u_{n+2} \geqslant u_{n+1}.$$

En raisonnant par récurrence, on peut alors montrer la propriété suivante :

Proposition 1.3.1 *Si f est une fonction croissante, alors toute suite récurrente (u_n) associée à f est monotone et le sens de variation de (u_n) est donné par le signe de la différence $u_1 - u_0$.*

Exercice : à faire.

ATTENTION : la croissance de f impose la *monotonie* de (u_n), mais de détermine pas le sens de variation de (u_n).

Exemple : déterminer le sens de variation des suites suivantes :

$$(u_n) : \begin{cases} u_0 = \frac{1}{2} \\ u_{n+1} = u_n^3 \end{cases} \quad ; \quad (v_n) : \begin{cases} v_0 = 2 \\ v_{n+1} = v_n^3 \end{cases}$$

Point fixe et intervalle stable. On appelle *point fixe* de f toute solutions x^* de l'équation $f(x) = x$ car on peut montrer que si (u_n) est une suite récurrente associée à f dont le point de départ est $u_0 = x^*$, alors (u_n) est constante égale à x^*.

(Exercice : à faire)

D'autre part, on peut montrer que les points fixes de f jouent un rôle fondamental dans l'évolution de toute suite récurrente (u_n) associée à f (même celle dont le point de départ u_0 n'est pas un point fixe x^* de f). En effet, si au départ $u_0 < x^*$, puisque f est croissante, on a

$$u_1 = f(u_0) \leqslant f(x^*) = x^*$$

Par récurrence, on peut alors montrer que $u_n \leqslant x^*$ pour tout n. Autrement dit, si $u_0 \leqslant x^*$, alors (u_n) est majorée par x^*.

On peut évidement montrer à l'inverse que si $u_0 > x^*$, alors $u_n \geqslant x^*$ pour tout $n \in \mathbb{N}$.

Formellement, on peut alors définir la notion *d'intervalles stables* pour une suite récurrente (u_n) associée à une fonction f.

Précisément, on appelle intervalle stable de (u_n) tout intervalle $[a, b]$ de $\mathbb{R}$ tel que si $u_0 \in [a, b]$ alors tous les termes de (u_n) sont également dans $[a, b]$.

Or les informations précédente se traduisent par le fait que les intervalles stables d'une suite récurrente (u_n) associée à une fonction f sont les intervalles définis par les points fixes de f. Autrement dit, si l'on note

$$x_1^* < x_2^* < \cdots < x_p^*$$

les solutions de l'équation $f(x) = x$, alors pour toute suite récurrente (u_n) associée à f, on a

$$u_0 \in [x_k^*, x_{k+1}^*] \ \Rightarrow \ u_n \in [x_k^*, x_{k+1}^*] \ \forall n \in \mathbb{N}.$$

On peut alors déterminer la nature de (u_n) dans certains cas particuliers en se basant uniquement sur les deux premiers termes u_0 et u_1 :

- Si $u_1 \geqslant u_0$ et $u_0 \leqslant x^*$, alors (u_n) est croissante et majorée par x^*. Elle est donc convergente.
- Si $u_1 \leqslant u_0$ et $u_0 \geqslant x^*$, alors (u_n) est décroissante et minorée par x^*. Elle est donc convergente.
- S'il existe deux points fixes x_1^* et x_2^* de f tels que $x_1^* \leqslant u_0 \leqslant x_2^*$, alors (u_n) est monotone et bornée, donc elle converge.

ATTENTION : les résultats précédents permettent de déterminer éventuellement la nature de (u_n) mais ne permettent pas de calculer la valeur de la limite éventuelle.

Cependant, en poussant l'étude de (u_n), on peut, après avoir établi sa nature convergente, en déduire la valeur de sa limite. En effet, si l'on sait que (u_n) converge et que l'on note ℓ sa limite, on a :

$$u_n \to \ell \qquad \text{et} \qquad u_{n+1} \to \ell.$$

En passant alors à la limite $n \to +\infty$ dans la relation

$$u_{n+1} = f(u_n),$$

on a, si f est continue

$$f(\ell) = \ell.$$

Autrement dit, si (u_n) est convergente, alors sa limite ℓ est nécessairement un point fixe de f. Cela restreint donc les limites possibles, et une étude un peu plus poussée doit nous permettre de conclure.

Exemple : soit $f : x \mapsto x^3$.

1. Déterminer les variations et les points fixes de f.
2. Soient
$$(u_n) : \begin{cases} u_0 = \frac{1}{2} \\ u_{n+1} = u_n^3 \end{cases} \qquad \text{et} \qquad (v_n) : \begin{cases} v_0 = 2 \\ v_{n+1} = v_n^3 \end{cases}$$
   (a) Montrer que (u_n) est minorée. En déduire qu'elle converge et donner sa limite.
   (b) Montrer que (v_n) ne peut être convergente.

Résumé : à partir des observations faites ci-dessus, on peut mettre en place à protocole d'étude pour toutes les suites récurrentes associées à une fonction f croissante basé sur l'étude du signe de la quantité $f(x) - x$: la première étape consiste à dresser le tableau de signe de $f(x) - x$. Cette étude produit en particulier la liste des points fixes $x_1^*, \ldots, x_p^*$ de f. Si l'on note alors

$$(u_n) : \begin{cases} u_0 \text{ donné} \\ u_{n+1} = f(u_n) \end{cases}$$

on peut énoncer les résultats suivants :

1. *Sens de variation.*
   - Toute suite (u_n) dont le point de départ u_0 est un point fixe de f est constante.
   - Toute suite (u_n) dont le point de départ u_0 est dans un intervalle où $f(x) - x \geqslant 0$ est croissante.
   - Toute suite (u_n) dont le point de départ u_0 est dans un intervalle où $f(x) - x \leqslant 0$ est décroissante.
2. *Convergence.*
   - Si la suite (monotone) (u_n) se rapproche d'un point fixe de f, alors elle converge vers ce point fixe.
   - Si (u_n) s'éloigne de tous les points fixes de f, alors elle diverge vers $\pm\infty$.

Exemple : déterminer le comportement de toutes les suites récurrentes vérifiant la relation $u_{n+1} = u_n^3$ en fonction du point de départ u_0.

Le cas des fonctions décroissantes. Dans toutes les études précédentes, on s'est restreint au cas où f est une fonction croissante. Or par une simple remarque, on va voir que l'on peut toujours se ramener à ce cas. En effet, si f est une fonction décroissante, en appliquant deux fois la fonction f, on obtient une fonction croissante.

(Exercice : à montrer).

Or si (u_n) est une suite récurrente associée à une fonction f décroissante, la suite récurrente définie par

$$(v_n) \ : \ \left\{ \begin{array}{l} v_0 = u_0 \\ v_{n+1} = f(f(v_n)) = f \circ f(v_n) \end{array} \right.$$

n'est autre que la suite extraite de (u_n) constituée des termes d'indices pairs :

$$\forall n \in \mathbb{N}, \quad v_n = u_{2n}.$$

(Exercice : à montrer par récurrence).

De même, la suite (w_n) définie par

$$(w_n) \ : \ \left\{ \begin{array}{l} w_0 = u_1 \\ w_{n+1} = f \circ f(w_n) \end{array} \right.$$

est la suite extraite de (u_n) constituée des termes d'indices impairs.

Or ces deux suites étant associées à la fonction croissante $g = f \circ f$, on peut les étudier à l'aide du protocole établi plus haut. D'autre part, puisqu'à elles deux, elles recouvrent

l'ensemble des termes de (u_n), leur étude permet la nature de (u_n) et sa limite éventuelle.

En poussant encore l'étude des ces suites, on peut montrer le résultat suivant :

Proposition 1.3.2 *Soit (u_n) une suite récurrente associée à une fonction f décroissante.*
Les suites extraites (u_{2n}) et (u_{2n+1}) sont monotones et de sens de variation différents. D'autre part :
- *Si la distance $|u_{2n} - u_{2n+1}|$ diminue, alors ces deux suites sont adjacentes et (u_n) converge vers l'unique point fixe de f contenu entre u_{2n} et u_{2n+1}.*
- *Si la distance $|u_{2n} - u_{2n+1}|$ augmente, alors (u_n) diverge.*

Exemples :

1. Déterminer le comportement de toute suite (u_n) associée à la fonction
$$f : x \longmapsto 1 + \frac{1}{x}$$
2. Déterminer le comportement de toute suite (v_n) associée à la fonction
$$g : x \longmapsto \frac{1}{x^2}$$

1.4 Vitesse de convergence d'une suite récurrente

D'autre part, les outils d'analyse permettent également de mesurer la vitesse de convergence d'une suite récurrente en étudiant en détails la fonction f associée.

1.4.1 Critère de convergence

L'un des intérêts majeurs des suites récurrentes est qu'elles permettent d'obtenir des valeurs approchées pour n'importe quel nombre réel. Il suffit pour cela de déterminer une fonction dont la valeur que l'on cherche est un point fixe.

Cependant, comme on l'a vu plus haut, il existe, pour une fonction f donnée, certains points fixes qui ne peuvent être approchés par une suite récurrente : les points fixes instables.

Pour approcher une valeur donnée, il faut donc construire une fonction pour laquelle la valeur cherchée est un point fixe *stable*.

Question : comment distinguer les points fixes stables des points fixes instables ?

Une première approche graphique permet de faire la différence : pour qu'un point fixe x^* soit stable, il faut que toute suite dont le point de départ $u_0 < x^*$ soit croissante et que

toute suite dont le point de départ $u_0 > x^*$ soit décroissante. Il faut donc que la quantité $f(x) - x$ soit

- positive si $x < x^*$,
- négative si $x > x^*$.

Il faut donc que la courbe de f coupe la première bissectrice "par au dessus". Autrement dit, autour de x^*, la croissance de f doit être moins rapide que celle de x.

D'un point de vue analytique, cela se traduit par le fait que la dérivée de f autour de x^* doit être inférieure à celle de $[x \mapsto x]$. On obtient alors le critère suivant :

Proposition 1.4.1 *Un point fixe x^* de f est*

- *stable si $f'(x^*) < 1$*
- *instable si $f'(x^*) > 1$.*

Note : si $f'(x^*) = 1$, on n'a pas de résultat, dans le sens ou tout peut se produire.

En conclusion, pour calculer une valeur approchée d'un nombre réel x^* donné, il faut construire une fonction f pour laquelle

1. $f(x^*) = x^*$
2. $f'(x^*) < 1$.

Exemples :

1. Déterminer la nature de chacun des points fixes de la fonction
$$f \,:\, x \mapsto \frac{1}{8}x^4 - x^3 + 2x^2.$$
2. Montrer que la fonction $f \,:\, x \mapsto \frac{1}{2}\left(x + \frac{2}{x}\right)$ permet de construire des suites récurrentes qui convergent vers $\sqrt{2}$.

1.4.2 Échelle de vitesses

Une fois que l'on sait construire une suite pour approcher un nombre réel donné, on s'intéresse à *la vitesse de convergence* de la suite que l'on a construit. Intuitivement, on souhaite pouvoir estimer le nombre de termes que l'on doit calculer pour obtenir une valeur approchée du nombre cherché à une précision donnée.

Commençons par établir une échelle de vitesses :

Soit (u_n) une suite numérique qui converge et soit ℓ sa limite.

1. On dit que la suite (u_n) converge *lentement* vers ℓ si
$$\text{il existe } A \text{ et } \alpha \ > 0 \text{ tels que } |u_n - \ell| \geqslant \frac{A}{n^\alpha}.$$
Intuitivement, cela signifie que (u_n) converge vers ℓ à une vitesse au plus polynomiale.
2. On dit que la suite (u_n) converge *géométriquement* vers ℓ si
$$\text{il existe } k \in]0,1[\text{ tel que } |u_n - \ell| \leqslant Ak^n$$
(i.e. la quantité $|u_n - \ell|$ est dominée par k^n et tend vers 0 à la vitesse de k^n).
3. On dit que la suite (u_n) converge avec une *convergence quadratique* vers ℓ si $|u_n - \ell|$ est dominée par k^{2^n} avec $0 < k < 1$.
4. Plus généralement, on dit que la suite (u_n) converge avec une *convergence d'ordre* $r > 1$ vers ℓ si $|u_n - \ell|$ est dominée par k^{r^n} avec $0 < k < 1$.

Question : comment mesurer la vitesse de convergence d'une suite donnée ?

Un indicateur important pour l'étude de la vitesse de convergence est le quotient
$$q_n = \frac{|u_{n+1} - \ell|}{|u_n - \ell|}$$
et plus précisément la limite $k = \lim q_n$.

On peut alors définir une nouvelle échelle de vitesse, globalement équivalente à la première :

1. Si $k = 1$, alors la convergence de (u_n) est *lente*.
2. Si $0 < k < 1$, alors la convergence est géométrique, de raison k.
3. Si $k = 0$, alors la convergence est rapide.

Remarque : si $k > 1$, le quotient q_n est supérieur à 1 à partir d'un certain rang. À partir de ce rang, la suite (u_n) s'éloigne de ℓ. Elle ne peut donc converger vers ℓ.

1.4.3 Application aux suites récurrente

Dans le cas particulier où (u_n) est une suite récurrente vérifiant une relation de la forme $u_{n+1} = f(u_n)$, convergent vers un point fixe x^* de f, on constate rapidement que
$$\frac{|u_{n+1} - x^*|}{|u_n - x^*|} = \left|\frac{f(u_n) - f(x^*)}{u_n - x^*}\right| \longrightarrow |f'(x^*)|.$$

Ainsi, la valeur de $f'(x^*)$ déterminer la nature stable ou instable d'un point fixe, mais elle détermine également la vitesse de convergence des suites qui convergent vers x^* :

1. Si $|f'(x^*)| = 1$, la convergence vers x^* est lente.
2. Si $|f'(x^*)| = k \in]0, 1[$, alors la convergence est géométrique de raison k.
3. Si $|f'(x^*)| = 0$, la convergence est au moins quadratique.

1.5 Méthode de Newton

La méthode de Newton est une méthode basée sur les suites récurrentes permettant de résoudre de façon approchée une équation de la forme $\varphi(x) = 0$.

Le principe analytique est de construire une suite récurrente qui converge rapidement vers la solution cherchée.

Newton est parti d'une constatation géométrique : si l'on se place en un point x_n de l'axe des abscisses et si la fonction φ n'a pas de variations trop brusques entre x_n et la racine x^* cherchée, alors le point d'intersection entre la tangente à $\mathcal{C}_f$ en x_n est l'axe des abscisses est significativement plus près de x^* que x_n.

En posant cette construction d'un point de vue analytique, on construit une suite récurrente (x_n) définie par régie par la relation de récurrence

$$x_{n+1} = x_n - \frac{\varphi(x_n)}{\varphi'(x_n)}.$$

La fonction f associée à cette suite récurrente est donc

$$f \,:\, x \mapsto x - \frac{\varphi(x)}{\varphi'(x)}$$

On peut montrer rapidement que tout point fixe de f est bien une solution de $\varphi(x) = 0$. De plus,

$$f'(x) = \frac{\varphi(x)\varphi''(x)}{\varphi'(x)^2}$$

donc $f'(x^*) = 0$. Toute suite convergente associée à cette fonction converge donc très rapidement.

Exemple : construire une suite récurrente qui converge rapidement vers la racine positive de $\varphi(x) = x^2 - 2$.

Chapitre 2

Espaces vectoriels

Introduction

L'algèbre linéaire rassemble l'ensemble des outils permettant d'étudier, de manipuler et d'utiliser la notion de vecteur.

L'utilisation la plus courant des vecteurs se fait en géométrie car la notion de repère et coordonnées permet de représenter les différentes positions du plan ou de l'espace sous forme numérique.

En effet, à l'aide de la notion de coordonnées, on peut représenter des points, des contours, des volumes, mais également des positions relatives, des déplacements...

Exemples :
- Si l'on se place dans un repère orthonormé du plan ou de l'espace, on peut définir *le produit scalaire* qui permet de déterminer si deux vecteurs donnés sont orthogonaux ou non à partir de leurs coordonnées.
- De même, *le produit vectoriel* permet de construire, à partir de deux vecteurs, un troisième vecteur orthogonal aux deux premiers, toujours en s'aidant des coordonnées.

Nous verrons que l'algèbre linéaire offre de nombreux outil de ce type (et d'autre bien plus élaborés) pour l'étude du plan ou de l'espace basé sur méthodes analytiques faisant intervenir les coordonnées.

Nous verrons en particulier que les matrices jouent un rôle fondamental en algèbre linéaire et qu'elles permettent en particulier de ramener la plupart des problèmes à la résolution d'un système linéaire.

Cependant, le développement et l'utilisation de ces outils est basée sur une théorisation

poussée de les notions de vecteurs et de coordonnées.

Pour bien en comprendre le fonctionnement et les différentes applications, il est donc nécessaire de connaître et comprendre les détails toutes ces notions.

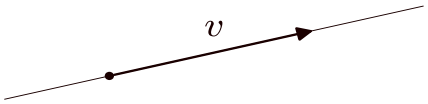
Après quelques rappels sur les vecteurs, nous verrons comment mettre en place le formalisme de l'algèbre linéaire.

Nous verrons ensuite quelles utilisation on peut faire de ce formalisme non seulement en algèbre linéaire, mais également dans de nombreux autres domaines de l'analyse et de l'algèbre.

2.1 Le plan et l'espace

2.1.1 Qu'est-ce qu'un vecteur ?

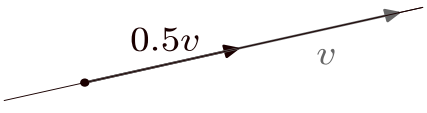
Dans le plan ou dans l'espace, un vecteur, c'est :
- un point de départ,
- une direction,
- un sens,
- une longueur.



Et il existe plusieurs opération que l'on peut faire à partir des vecteurs du plan. Ainsi, étant donné un vecteur du plan (ou de l'espace), on peut

- l'agrandir en le multipliant par une constante > 1,

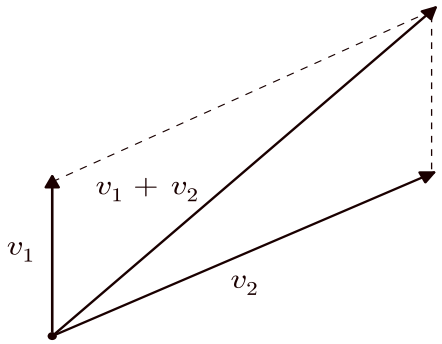
- le retressir en le multipliant par une constante < 1,



- le retourner en le multipliant par -1.

(Rappel : deux vecteurs de la forme $\overrightarrow{v}$ et $k.\overrightarrow{v}$ sont dits colinéaires.)

D'autre part, étant donnés deux vecteurs du plan (ou de l'espace), on peut les additionner pour construire un troisième vecteur du plan, en les mettant bout à bout.

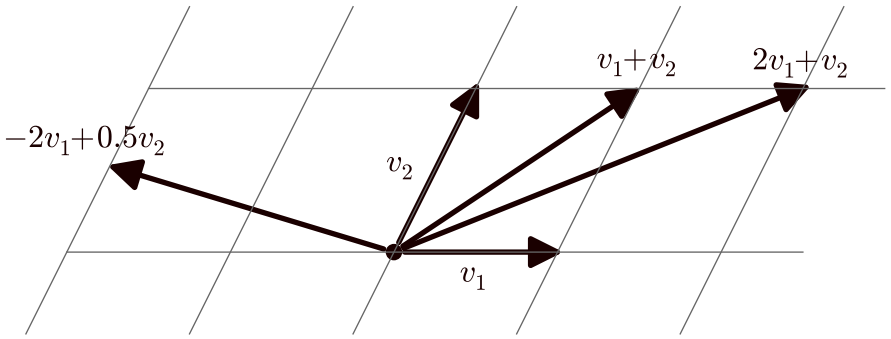


Ces deux opérations ont non seulement un sens (contrairement à la multiplication de deux vecteurs, par exemple), mais elles permettent également de mettre en place la notion de repère et de coordonnées.

2.1.2 Repères et coordonnées

Dans le plan, si l'on se donne un point d'origine et deux vecteurs $\overrightarrow{v_1}$ et $\overrightarrow{v_2}$ en prenant soin de ne pas les choisir colinéaires, on peut représenter n'importe quel vecteur du plan à l'aide de ces deux vecteurs.

Exemples : Exprimer en fonction de $\overrightarrow{v_1}$ et $\overrightarrow{v_2}$ les trois vecteurs ci dessous :



L'ensemble $\mathcal{R} = (0; \overrightarrow{v_1}, \overrightarrow{v_2})$ est alors un repère du plan et pour tout vecteur $\overrightarrow{u}$ du plan, il existe des réels x et y tels que

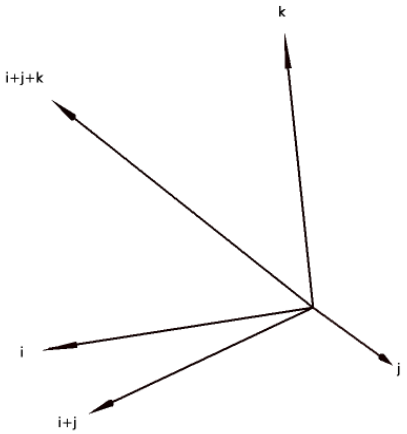
$$\overrightarrow{u} = x.\overrightarrow{v_1} + y.\overrightarrow{v_2}.$$

Les réels (x, y) sont les coordonnées de $\overrightarrow{u}$ dans le repère $\mathcal{R}$. On note $\overrightarrow{u} = \begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{R}}$.

Si l'on ajoute un vecteur $\overrightarrow{k}$ au repère, en le prenant hors du plan, on peut alors écrire tout vecteur de l'espace en fonction de $\overrightarrow{i}$, $\overrightarrow{j}$ et $\overrightarrow{k}$. Autrement dit, pour tout vecteur $\overrightarrow{v}$ de l'espace, il existe des nombres (x, y, z) tels que

$$\overrightarrow{v} = x.\overrightarrow{i} + y.\overrightarrow{j} + z.\overrightarrow{k}$$

Ces nombres x, y et z sont les coordonnées de $\overrightarrow{v}$ dans le repère $(0; \overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$



Note : dans le plan, il faut prendre soin de choisir des vecteurs non colinéaires. En effet, si l'on choisit deux vecteurs $\overrightarrow{i}$ et $\overrightarrow{j}$ colinéaires, les seuls vecteurs du plan que l'on pourra écrire en fonction de $\overrightarrow{i}$ et $\overrightarrow{j}$ sont les vecteur de la droite les portant tous les deux.

De même, il faut bien choisir les vecteurs de l'espace si l'on veut pouvoir traiter tous les vecteurs de l'espace. De façon générale, ils ne doivent pas être co-planaires. En effet, s'ils sont dans un même plan, les seuls vecteurs que l'on pourra exprimer en fonction de $\overrightarrow{i}$, $\overrightarrow{j}$ et $\overrightarrow{k}$ sont les vecteur du plan les rassemblant...

On verra que cette idée peut être formaliser et que le calcul matriciel donne un outil efficace pour faire le tri dans les différentes familles de vecteurs.

2.1.3 L'idée d'espace vectoriel

Les notions que l'on vient de voir, notamment l'idée de repère et de coordonnées, permettent de définir des positions dans l'espace et dans le plan. En effet, une fois que l'on s'est fixé un point d'origine O, on peut représenter tout point M du plan ou de l'espace par le vecteur $\overrightarrow{OM}$. Si l'on complète de plus le point O en un repère, on peut alors représenter tout point M par ses coordonnées dans ce repère. Ces coordonnées nous permettent alors de transformer tout problème de géométrie du plan ou de l'espace (position relatives de plusieurs points, déplacement d'un point, d'une figure...) en un problème de calcul.

Par exemple, si $\overrightarrow{OM}$ est un point du plan, de coordonnées $\begin{pmatrix} x \\ y \end{pmatrix}_{\mathcal{R}}$ (dans un repère $\mathcal{R} = (O; \overrightarrow{i}, \overrightarrow{j})$ fixé), le vecteur $2.\overrightarrow{OM}$ aura pour coordonnées $\begin{pmatrix} 2x \\ 2y \end{pmatrix}_{\mathcal{R}}$. De même, si $\overrightarrow{OM_1}$ et $\overrightarrow{OM_2}$ ont pour coordonnées respectives $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}_{\mathcal{R}}$ et $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}_{\mathcal{R}}$, le vecteur $\overrightarrow{OM_1} + \overrightarrow{OM_2}$ aura pour coordonnées $\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}_{\mathcal{R}}$.

Un autre exemple important : si l'on munit l'espace d'un repère $\mathcal{R} = (O; \overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k})$, alors tous les vecteurs de l'espace appartenant à un même plan passant par l'origine vérifient une même équation du type $ax + by + cz = 0$, où a, b, c sont des réels fixés et x, y, z sont les coordonnées dans $\mathcal{R}$.

Le premier objectif de ce cours est de formaliser et étudier les objets que l'on vient de voir et d'établir certaine propriétés générales. On appellera alors espace vectoriel tout ensemble vérifiant ces propriétés. Cela nous permettra notamment de voir certaines classes d'objet (comme l'ensemble des polynômes, ou celui des matrices) comme des espaces vectoriels. Et en étudiant en détails la notion d'espace vectoriels, on obtiendra des outils permettant de travailler de la même façon avec des point du plan ou de l'espace, des polynômes ou des matrices...

2.2 Définitions et Exemples

2.2.1 Loi de composition interne

Commençons par formaliser l'addition de deux vecteur. De façon formelle, l'addition est une opération qui combine deux éléments du plan (ou de l'espace) pour en construire un troisième. C'est ce que l'on appelle une loi de composition interne (lci). On peut en outre remarquer que l'addition de deux vecteurs que l'on vient de voir possède quelques propriétés élémentaires :

- L'addition des vecteurs est commutative : pour tous vecteurs $\overrightarrow{u}$ et $\overrightarrow{v}$, on a
$$\overrightarrow{u} + \overrightarrow{v} = \overrightarrow{v} + \overrightarrow{u}.$$
- L'addition des vecteurs est associative : pour tous vecteurs $\overrightarrow{u}$, $\overrightarrow{v}$ et $\overrightarrow{w}$, on a
$$(\overrightarrow{u} + \overrightarrow{v}) + \overrightarrow{w} = \overrightarrow{u} + (\overrightarrow{v} + \overrightarrow{w}).$$
- L'addition des vecteurs possède un élément neutre. C'est-à-dire un vecteur qui ne fait rien. Il s'agit bien sur du point O (ou du vecteur nul $\overrightarrow{0}$).
$$\overrightarrow{u} + \overrightarrow{0} = \overrightarrow{u}.$$
- Tout vecteur $\overrightarrow{v}$ possède un inverse pour la l'addition, c'est-à-dire un autre vecteur tel que leur somme soit égale au vecteur nul. C'est le vecteur $-\overrightarrow{v}$.

2.2.2 Multiplication externe

De la même façon, on peut formaliser la multiplication d'un vecteur par un réel $\overrightarrow{v} \mapsto k.\overrightarrow{v}$. Cette opération est une multiplication externe, puisque l'on utilise un élément qui

n'est pas dans notre ensemble (le réel k). D'autre part, la multiplication externe que l'on a vu possède elle aussi quelques propriétés :

- La multiplication externe est distributive par rapport à l'addition des vecteurs :

$$k.(\overrightarrow{u} + \overrightarrow{v}) = k.\overrightarrow{u} + k.\overrightarrow{v}.$$

- La multiplication externe est distributive par rapport à l'addition des réels :

$$(k + \ell).\overrightarrow{u} = k.\overrightarrow{u} + \ell.\overrightarrow{u}.$$

- La multiplication externe est associative :

$$(k\ell).\overrightarrow{u} = k.(\ell.\overrightarrow{u}).$$

- La multiplication externe possède un élément neutre : c'est le réel 1 :

$$1.\overrightarrow{v} = \overrightarrow{v}.$$

2.2.3 Espace vectoriel

Ce sont les opérations que l'on peut définir sur un ensemble et les propriétés de ces opérations qui font de cet ensemble un espace vectoriel. Précisément,

Définition 2.2.1 *Soit E un ensemble. Cet ensemble est un espace vectoriel si l'on peut le munir*
- *d'une lci commutative, associative et possédant un élément neutre et telle que tout élément de E soit inversible pour cette lci,*
- *d'une multiplication externe par des réels qui soit distributive (aux sens où on l'a vu plus haut), associative et possédant également un élément neutre.*

On appellera vecteur tout élément d'un espace vectoriel.

Si l'on note $+$ et $.$ la lci et la multiplication d'un espace vectoriel E, on appellera combinaison linéaire toute opération sur les vecteurs de E combinant les deux opérations, i.e. toute opération de la forme

$$k_1.\overrightarrow{v_1} + k_2.\overrightarrow{v_2} + \ldots + k_n.\overrightarrow{v_n}$$

où les $\overrightarrow{v_i}$ sont des vecteurs de E et les k_i sont des réels.

Dans la suite, on utilisera souvent les combinaisons linéaires pour construire un vecteur à partir d'une famille de vecteurs donnée.

Exemples d'espaces vectoriels.

1. Le plan ou l'espace (muni d'un point d'origine O) est un espace vectoriel.
2. La droite munie d'un point d'origine est un espace vectoriel.
3. L'ensemble des polynômes muni des opérations naturelles d'addition et de multiplication par un réel est un espace vectoriel.
4. L'ensemble des matrices d'une même taille, muni des opérations naturelles d'addition et de multiplication par un réel est un espace vectoriel.
5. Considérons l'ensemble $\mathbb{R}^2$ des couples de réels.

   $$\mathbb{R}^2 = \{(x, y),\ x \in \mathbb{R},\ y \in \mathbb{R}\}.$$

   On peut munir $\mathbb{R}^2$ d'une addition

   $$(x, y) + (x', y') = (x + x', y + y')$$

   et d'une multiplication par un réel

   $$k.(x, y) = (kx, ky).$$

   L'ensemble $\mathbb{R}^2$ muni de ces deux lois est un espace vectoriel.
6. De façon plus général, pour un entier n donné, l'ensemble $\mathbb{R}^n$ des n-uplets à coordonnées dans $\mathbb{R}$, muni des mêmes opérations (coordonnée par coordonnée) est un espace vectoriel.
7. L'ensemble $\mathcal{C}$ des couples $(x, y) \in \mathbb{R}^2$ tels que $0 \leqslant x, y \leqslant 1$ (noté $[0, 1]^2$) n'est pas un espace vectoriel. En effet, bien que l'on puisse définir sur $\mathcal{C}$ les opérations "+" et "." nécessaires (les mêmes que sur $\mathbb{R}^2$ tout entier), plusieurs des propriétés d'un espace vectoriel ne sont pas vérifiées :
   (a) La loi + n'est pas une lci : bien que $\left(\frac{1}{2}, \frac{1}{2}\right)$ et $(1, 1)$ soient $\mathcal{C}$, leur somme

   $$\left(\frac{1}{2}, \frac{1}{2}\right) + (1, 0) = \left(\frac{3}{2}, \frac{1}{2}\right)$$

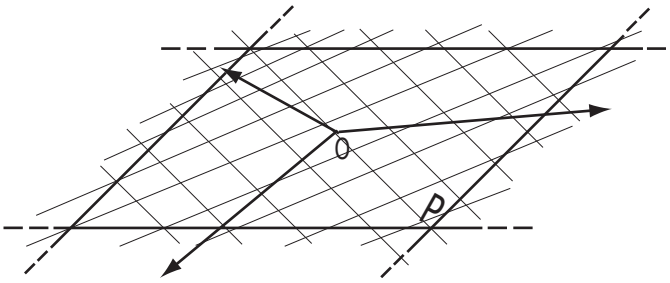
   n'est pas dedans.
   (b) Notre ensemble contient bien un élément neutre pour "+" (c'est $(0, 0)$), mais certains éléments de $[0, 1]^2$ n'ont pas d'inverse pour la loi +, par exemple $\left(\frac{1}{2}, 0\right)$.
   (c) Notre ensemble n'est pas stable par la multiplication. Par exemple

   $$4 \times \left(\frac{1}{2}, \frac{1}{2}\right) = (2, 2) \notin [0, 1]^2.$$

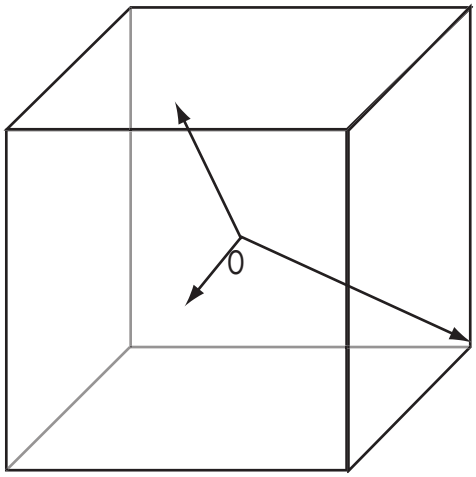
2.2.4 Sous espaces vectoriels

Plaçons nous maintenant dans un espace vectoriel E, muni d'une lci "+" et d'une multiplication " . " (par exemple, l'espace dans lequel on fixe un point O) et étudions un peu ses sous-ensembles.

On peut par exemple considérer l'ensemble des vecteurs de l'espace qui sont dans un même plan $\mathcal{P}$,



ou l'ensemble des vecteurs de l'espace qui sont contenus dans un cube centré en O.



Parmi ces sous ensemble, il en est certains que l'on peut munir d'une structure d'espace vectoriels. C'est ce qu'on appelle *les sous-espaces vectoriels* de E.

Pour qu'un sous ensemble de V soit un sous espace vectoriel de E, il faut donc qu'il vérifie certaines propriétés.

Théorème 2.2.2 *Étant donné un espace vectoriel E, un sous ensemble $\mathcal{S}$ de E est* un sous espace vectoriel *de E si et seulement si*

1. *L'élément neutre $\overrightarrow{0}_E$ de E est dans $\mathcal{S}$.*
2. *$\mathcal{S}$ est fermé pour la loi + :*

$$\forall \overrightarrow{u}, \overrightarrow{v} \in \mathcal{S}, \quad \overrightarrow{u} + \overrightarrow{v} \in \mathcal{S}$$

3. *$\mathcal{S}$ est fermé pour la loi . :*

$$\forall k \in \mathbb{R}, \forall \overrightarrow{u} \in \mathcal{S}, \quad k.\overrightarrow{u} \in \mathcal{S}$$

Les deux derniers points peuvent être rassemblés dans la propriété suivante : pour tous $k, \ell \in \mathbb{R}$ et pour tous $\overrightarrow{u}, \overrightarrow{v} \in \mathcal{S}$, on a

$$k.\overrightarrow{u} + \ell.\overrightarrow{v} \in \mathcal{S}.$$

Exemples de SEV

1. Dans le plan muni d'un point d'origine O, l'ensemble des vecteurs colinéaires à un vecteur donné est un SEV. En effet, soit $\overrightarrow{v}$ un vecteur du plan. L'ensemble des vecteurs colinéaires à $\overrightarrow{v}$ est

$$V = \{k.\overrightarrow{v},\ k \in \mathbb{R}\}.$$

Pour montrer que V est un SEV du plan, on doit vérifier que V
– contient le vecteur nul,
– est stable par addition,
– est stable par multiplication extérieure.

Or
– $\overrightarrow{0} = 0.\overrightarrow{v}$. Le vecteur nul est donc bien de la forme $k.\overrightarrow{v}$.
– Si $\overrightarrow{u_1}$ et $\overrightarrow{u_2}$ sont deux vecteurs de V, ils s'écrivent

$$\overrightarrow{u_1} = k_1.\overrightarrow{v} \qquad \text{et} \qquad \overrightarrow{u_2} = k_2.\overrightarrow{v}.$$

La somme $\overrightarrow{u_1} + \overrightarrow{u_2}$ s'écrit alors

$$\overrightarrow{u_1} + \overrightarrow{u_2} = k_1.\overrightarrow{v} + k_2.\overrightarrow{v} = (k_1 + k_2).\overrightarrow{v}.$$

Cette somme est donc de la forme $k.\overrightarrow{v}$ avec $k = k_1 + k_2 \in \mathbb{R}$ et V est stable par addition.

– Si $\overrightarrow{u} = \alpha.\overrightarrow{v}$ est un vecteur de V, pour tout $\beta \in \mathbb{R}$, on a

$$\beta.\overrightarrow{u} = \beta.(\alpha.\overrightarrow{v}) = (\beta.\alpha).\overrightarrow{v}.$$

$\beta.\overrightarrow{u}$ est donc bien de la forme $k.\overrightarrow{v}$ avec $k = \beta.\alpha$. C'est donc un vecteur de V et V est stable par multiplication extérieure.

2. Dans l'espace, tout plan contenant O est un SEV, toute droite passant par O est un SEV, le point O est un SEV.
3. Le cube de l'espace que l'on vient de voir n'est pas un SEV car il n'est pas fermé pour les opérations + et .

4. Dans $\mathbb{R}^3$ l'ensemble des vecteurs de la forme $(a, b, 0)$ pour a et b parcourant $\mathbb{R}$ est un sous espace vectoriel de $\mathbb{R}^3$.
   En effet, Soit $\mathcal{S} = \{(a, b, 0),\ a, b \in \mathbb{R}\}$. Le vecteur nul $(0, 0, 0)$ est bien un élément de $\mathcal{S}$. De plus, soient $k, \ell \in \mathbb{R}$ et soient $(a, b, 0)$ et $(a', b', 0)$ deux vecteurs de $\mathcal{S}$. Alors

$$\begin{aligned} k.(a, b, 0) + \ell.(a', b', 0) &= (ka, kb, 0) + (\ell a', \ell b', 0) \\ &= (ka + \ell a', kb + \ell b', 0) \end{aligned}$$

   est bien de la forme $(c, d, 0)$.

   Par contre, l'ensemble $\mathcal{S}' = \{(a, b, 1),\ a, b \in \mathbb{R}\}$ n'est pas un SEV de $\mathbb{R}^3$. Pourquoi ?

5. L'ensemble des solutions d'une équation linéaire homogène à 2 (resp. trois variables) est un SEV de $\mathbb{R}^2$ (resp. $\mathbb{R}^3$). En effet, une équation linéaire homogène en deux variables (x, y) est une équation de la forme $ax + by = 0$, $a, b \in \mathbb{R}$ connus. Soit, par exemple

$$(E) : x + y = 0.$$

   L'ensemble des solutions de (E) est donc

$$S = \{(x, y) \in \mathbb{R}^2 \ / \ x + y = 0\}.$$

   Or
   – Puisque $0 + 0 = 0$, le vecteur $(0, 0)$ fait bien partie de l'ensemble des solutions.

   – Si $\overrightarrow{v_1} = (x_1, y_1)$ et $\overrightarrow{v_2} = (x_2, y_2)$ sont deux solutions de (E) (i.e. $x_1 + y_1 = 0$ et $x_2 + y_2 = 0$), la somme $\overrightarrow{v_1} + \overrightarrow{v_2}$ est encore une solution. En effet, calculons les coordonnées de $\overrightarrow{v_1} + \overrightarrow{v_2}$:

$$\overrightarrow{v_1} + \overrightarrow{v_2} = (x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2).$$

     Ce vecteur est dans S si la somme de ses coordonnées est nulle. Or

$$(x_1 + x_2) + (y_1 + y_2) = (x_1 + y_1) + (x_2 + y_2) = 0 + 0 = 0$$

     car $\overrightarrow{v_1}$ et $\overrightarrow{v_2}$ sont dans S. S est donc stable par addition.

   – Si $\overrightarrow{v_1} = (x_1, y_1)$ est dans S, pour tout $k \in \mathbb{R}$, on a :

$$k.\overrightarrow{v_1} = k.(x_1, y_1) = (k.x_1, k.y_1)$$

     et la somme de ces coordonnées vaut

$$k.x_1 + k.y_1 = k.(x_1 + y_1) = k.0 = 0.$$

     le produit $k.\overrightarrow{v_1}$ est encore dans (S).

D'un point de vue géométrique, cet exemple est à rapprocher du premier exemple. En effet, dans les deux cas, l'ensemble étudié peut être représenté par une droite du plan. Précisément, il est possible de montrer que l'ensemble V que l'on vient d'étudier correspond à l'ensemble des vecteurs du plan qui sont colinéaires à un vecteur $\overrightarrow{v}$ donné. Il est de plus possible de déterminer ce vecteur à partir de l'équation : soit $\overrightarrow{u} = (x, y)$ un vecteur de V. On a donc

$$x + y = 0 \Longleftrightarrow y = -x.$$

Autrement dit, les coordonnées de $\overrightarrow{u}$ sont de la forme $(x, -x) = x.(1, -1)$ pour $x \in \mathbb{R}$. L'ensemble V est donc l'ensemble des vecteurs colinéaires au vecteur $\overrightarrow{v} = (1, -1)$.

6. L'ensemble des solutions d'un système linéaire homogène 3×3 est un SEV de l'espace vectoriel $\mathbb{R}^3$.

   Rappelons qu'un système linéaire 3×3 est un système de la forme

$$\left\{ \begin{array}{rcl} a_1x + b_1y + c_1z & = & 0 \\ a_2x + b_2y + c_2z & = & 0 \\ a_3x + b_3y + c_3z & = & 0 \end{array} \right. \quad \text{ou} \quad \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

   Exemple :

$$(E) : \left\{ \begin{array}{rcl} 2x - 2y + 2z & = & 0 \\ -x + y - z & = & 0 \\ x + y + z & = & 0 \end{array} \right.$$

   Les solutions de ce système sont $\left\{ \begin{array}{l} y = 0 \\ z = -x \end{array} \right.$. Autrement dit, tous les triplets $(x, y, z) \in \mathbb{R}^3$ qui s'écrivent $(x, 0, -x)$ pour une certaine valeur de x. Si les inconnues x, y et z de notre système représentent les coordonnées de certains points dans un repère $\mathcal{R}$ de l'espace, l'ensemble des points solution est donc l'ensemble des points de la forme

$$\begin{pmatrix} x \\ 0 \\ -x \end{pmatrix}_{\mathcal{R}} = x. \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}_{\mathcal{R}}.$$

   C'est donc la droite portée par le vecteur $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}_{\mathcal{R}}$.

   On peut en outre montrer de façon générale que l'ensemble des solutions d'un système linéaire homogène est un SEV en montrant qu'un tel ensemble contient toujours $\overrightarrow{0}$ et qu'il est stable par combinaisons linéaires.

2.2.5 Sommes de sous espaces vectoriels

Soit E un espace vectoriel et soient $\mathcal{W}_1$ et $\mathcal{W}_2$ deux sous espaces vectoriels de E. À partir de ces deux sous espaces et à l'aide de la loi "+" que l'on a défini sur E, on peut construire un nouveau sous ensemble de E : la somme $\mathcal{W}_1 + \mathcal{W}_2$ contenant tous les vecteurs $v \in E$ qui s'écrivent comme la somme d'un élément de $\mathcal{W}_1$ et 'un élément de $\mathcal{W}_2$:

$$\mathcal{W}_1 + \mathcal{W}_2 = \{v = w_1 + w_2,\ w_1 \in \mathcal{W}_1,\ w_2 \in \mathcal{W}_2\}.$$

Proposition 2.2.3 *La somme de deux sous espaces vectoriels est encore un sous espace vectoriel.*

Preuve : soit E un espace vectoriel, et soient $\mathcal{W}_1$ et $\mathcal{W}_2$ deux sous espaces vectoriels de E. Pour montrer que $\mathcal{W}_1 + \mathcal{W}_2$ est un sous espace vectoriel, on utilise le théorème 2.2.2.

1. Comme $\mathcal{W}_1$ et $\mathcal{W}_2$ sont des espaces vectoriels, ils contiennent le vecteur 0_E. Donc la somme contient également $0_E = 0_E + 0_E$.
2. Soient maintenant v et v' deux vecteurs de $\mathcal{W}_1 + \mathcal{W}_2$ et soient k et ℓ deux réels. Pour montrer que la somme $\mathcal{W}_1 + \mathcal{W}_2$ est un espace vectoriel, il faut montrer que $k.w_1 + \ell.w_2$ est encore dans $\mathcal{W}_1 + \mathcal{W}_2$. Or par définition, il existe des vecteurs $w1, w'_1, w_2$ et w'_2 tels que

$$\begin{aligned} v &= w_1 + w_2 \\ v' &= w'_1 + w'_2 \end{aligned}$$

Ainsi,

$$\begin{aligned} k.v_1 + \ell.v_2 &= k.(w_1 + w_2) + \ell.(w'_1 + w'_2) \\ &= (k.w_1 + \ell.w'_1) + (k.w_2 + \ell.w'_2) \end{aligned}$$

Comme $\mathcal{W}_1$ et $\mathcal{W}_2$ sont des sous espaces vectoriels, on a

$$k.w_1 + \ell.w'_1 \in \mathcal{W}_1 \quad \text{et} \quad k.w_2 + \ell.w'_2 \in \mathcal{W}_2.$$

Le vecteur $k.v + \ell.v'$ est donc bien un élément de $\mathcal{W}_1 + \mathcal{W}_2$.

En conclusion $\mathcal{W}_1 + \mathcal{W}_2$ est un sous ensemble non vide de E, fermé pour les lois "+" et " . " . C'est donc un sous espace vectoriel de E □

Définition 2.2.4 *La somme $\mathcal{W}_1 + \mathcal{W}_2$ sera dite* directe *(et notée $\mathcal{W}_1 \oplus \mathcal{W}_2$) si l'on a*

$$\mathcal{W}_1 \cap \mathcal{W}_2 = \{0_E\}.$$

Remarquons que l'intersection de deux sous espaces vectoriels contient toujours le vecteur nul. La somme est donc directe si l'intersection est *minimale*.

Définition 2.2.5 *Si deux SEV $\mathcal{W}_1$ et $\mathcal{W}_2$ sont en somme directe et vérifient*

$$E = \mathcal{W}_1 \oplus \mathcal{W}_2$$

on dit que $\mathcal{W}_1$ et $\mathcal{W}_2$ sont supplémentaires dans E.

Proposition 2.2.6 *Si $\mathcal{W}_1$ et $\mathcal{W}_2$ sont supplémentaires dans E, alors tout vecteur $v \in E$ s'écrit* de façon unique *sous la forme*

$$v = w_1 + w_2, \quad w_1 \in \mathcal{W}_1,\ w_2 \in \mathcal{W}_2.$$

Exemples

1. Dans l'espace, la somme de deux droite non colinéaires est directe et recouvre le plan contenant ces deux droites.
2. Dans le plan, deux droites non colinéaires sont supplémentaires.
3. Dans l'espace, un plan et une droite extérieure à ce plan sont supplémentaires.
4. Dans l'espace, deux plan ne sont jamais en somme directe.

2.3 Familles de vecteurs

Dans cette section, nous allons revenir en détails sur la notion de repère que l'on a vu plus haut. Intuitivement, il s'agit d'étudier le fait que dans un espace vectoriel, il est possible de trouver une famille de vecteurs (souvent en nombre fini) qui nous permette de reconstruire tous les vecteurs de l'espace via les combinaisons linéaires.

Pour définir proprement la notion de repère, il faut étudier en détails les différents types de familles que l'on construire en "piochant" quelques vecteur dans même espace vectoriel.

Enfin, une fois établie la définition d'un repère, nous verrons comment définir proprement la notion de dimension (qui n'est pour l'instant qu'une notion intuitive).

2.3.1 Familles génératrices

Soient $\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}$ des vecteurs de E. On peut construire à l'aide des v_i un sous espace vectoriel de E particulier : le sous espace engendré par les $\overrightarrow{v_i}$.

Définition 2.3.1 *1. Soit E un espace vectoriel et soient $\overrightarrow{v_1}, \ldots, \overrightarrow{v_n}$ des éléments de E. Le sous espace engendré par les $\overrightarrow{v_i}$ est l'ensemble des vecteurs obtenus par combinaisons linéaires des $\overrightarrow{v_i}$.*

$$V(\overrightarrow{v_1}, \ldots, \overrightarrow{v_n}) = \{k_1.\overrightarrow{v_1} + k_2.\overrightarrow{v_2} + \ldots + k_n.\overrightarrow{v_n},\ k_i \in \mathbb{R}\}.$$

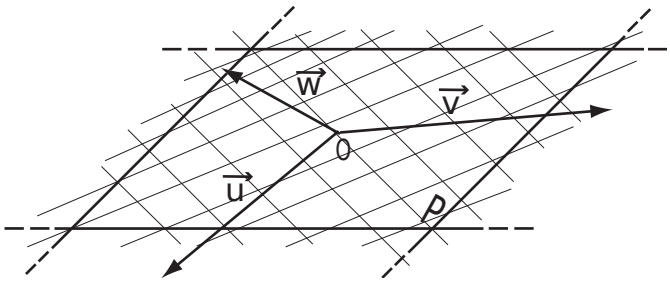
2. *Soit E un espace vectoriel. S'il existe une famille $\mathcal{F} = \{\overrightarrow{v_1}, \ldots, \overrightarrow{v_n}\}$ telle que $E = V(\overrightarrow{v_1}, \ldots, \overrightarrow{v_n})$, on dira que $\mathcal{F}$ est une famille génératrice de E.*

On sera souvent amenés à chercher une famille génératrice d'un espace donné. Ainsi, en reprenant l'exemple du plan, on voit que pour obtenir une famille génératrice du plan, il nous faut au moins deux vecteurs non colinéaires. Intuitivement, il nous faut une famille donnant au moins deux directions différentes. Concernant l'espace, on voit là qu'il nous faut au moins trois directions. De façon générale, pour obtenir une famille génératrice pour un espace vectoriel donné, il faut s'assurer que chacune des directions de notre espace vectoriel est représentée.

Exemples

1. Deux vecteurs non linéaires appartenant à un même plan engendrent le plan tout entier.
2. N'importe quel vecteur (non nul !) d'une droite engendre la droite toute entière.
3. Dans l'espace, la famille $\{\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}\}$ donnée par le dessin ci-dessous n'engendre pas l'espace tout entier. A l'aide des combinaisons linéaires, on ne pourra pas construire de vecteur sortant du plan $\mathcal{P}$.

   Par contre, la famille $\{\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}\}$ engendre le plan $\mathcal{P}$ car tout vecteur de $\mathcal{P}$ peut s'écrire comme une combinaison linéaire des vecteurs $\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}$.



4. Si l'on reprend l'exemple précédent dans $\mathbb{R}^3$, avec $\mathcal{S} = \{(a, b, 0),\ a, b \in \mathbb{R}\}$, on peut voir que $\mathcal{S}$ est le sous espace vectoriel de $\mathbb{R}^3$ engendré par les vecteurs $\overrightarrow{v_1} = (1, 0, 0)$ et $\overrightarrow{v_2} = (0, 1, 0)$.

   Preuve : pour montrer que $\mathcal{S} = V(\overrightarrow{v_1}, \overrightarrow{v_2})$, il faut montrer d'une part que toute combinaison linéaire de $\overrightarrow{v_1}$ et $\overrightarrow{v_2}$ est un vecteur de $\mathcal{S}$ (i.e. $V(\overrightarrow{v_1}, \overrightarrow{v_2}) \subset \mathcal{S}$) et d'autre part que tout vecteur de $\mathcal{S}$ peut s'écrire comme une combinaison linéaire de $\overrightarrow{v_1}$ et $\overrightarrow{v_2}$ (i.e. $\mathcal{S} \subset V(\overrightarrow{v_1}, \overrightarrow{v_2})$).

   (a) $V(\overrightarrow{v_1}, \overrightarrow{v_2}) \subset \mathcal{S}$: soit $\overrightarrow{v}$ un vecteur de $V(\overrightarrow{v_1}, \overrightarrow{v_2})$. Par définition, v est une combinaison linéaire des vecteurs $\overrightarrow{v_1}$ et $\overrightarrow{v_2}$. Autrement dit, il existe des réels a et b tels que $\overrightarrow{v} = a.\overrightarrow{v_1} + b.\overrightarrow{v_2}$. Mais alors, d'après la définition des opérations + et .

que l'on a sur $\mathbb{R}^2$, on peut écrire

$$\begin{aligned} v = a.\overrightarrow{v_1} + b.\overrightarrow{v_2} &= a.(1,0,0) + b.(0,1,0) \\ &= (a,0,0) + (0,b,0) \\ &= (a,b,0). \end{aligned}$$

C'est donc bien un élément de $\mathcal{S}$.

(b) $\mathcal{S} \subset V(\overrightarrow{v_1}, \overrightarrow{v_2})$: soit $\overrightarrow{v}$ un vecteur de $\mathcal{S}$. Par définition, il existe des réels a et b tel que les coordonnées de $\overrightarrow{v}$ s'écrivent

$$\overrightarrow{v} = (a,b,0).$$

Mais alors, toujours d'après la définition des opérations + et . que l'on a sur $\mathbb{R}^2$, on peut écrire

$$\begin{aligned} \overrightarrow{v} = (a,b,0) &= (a,0,0) + (0,b,0) \\ &= a.(1,0,0) + b.(0,1,0) \\ &= a.\overrightarrow{v_1} + b.\overrightarrow{v_2}. \end{aligned}$$

Notre vecteur $\overrightarrow{v}$ s'écrit donc comme combinaison linéaires des vecteurs $\overrightarrow{v_1}$ et $\overrightarrow{v_2}$. C'est donc un vecteur de $V(\overrightarrow{v_1}, \overrightarrow{v_2})$. □

2.3.2 Familles libres

On vient de voir que pour construire une famille génératrice, il nous fallait nécessairement un nombre suffisant de vecteurs. Une fois que toutes les directions de notre espace sont représentées, tout vecteur que l'on ajoutera à notre famille génératrice n'apportera pas d'information supplémentaire. Cette redondance est ce que l'on appelle de la dépendance linéaire. Précisément,

Définition 2.3.2 *Une famille $\{u_1, \ldots, u_m\}$ de vecteurs de E sera dite* linéairement dépendante *(ou* liée*) si un vecteur peut s'écrire comme combinaison linéaire des autres.*

Une famille qui n'est pas liée sera dite libre *(ou* linéairement indépendante*).*

Intuitivement, une famille est libre si chaque vecteur apporte une direction supplémentaire.

Ainsi, si une famille $\{\overrightarrow{v_1}, \ldots, \overrightarrow{v_n}\}$ est linéairement dépendante, on peut écrire le vecteur $\overrightarrow{v_1}$ sous la forme

$$\overrightarrow{v_1} = k_2.\overrightarrow{v_2} + \ldots + \overrightarrow{v_n}.$$

Le vecteur $\overrightarrow{v_1}$ n'apporte donc pas de direction supplémentaire, et l'on a

$$V(\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_n}) = V(\overrightarrow{v_2}, \ldots, \overrightarrow{v_n}).$$

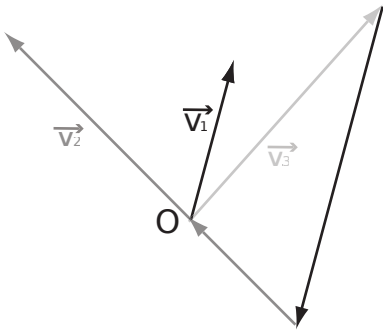
Parmi les famille génératrices d'un espace vectoriel donné, celles qui ne sont pas liées sont donc optimales (i.e. elles contiennent un nombre minimal de vecteurs). Quand on voudra construire une famille génératrice, on prendra donc soin d'en choisir une qui ne soit pas liée. Il est donc pratique de pouvoir déterminer si une famille donnée est liée ou non. Pour cela, on a une définition plus pratique de la dépendance linéaire :

Définition 2.3.3 *Une famille* $\{u_1, \ldots, u_m\}$ *de vecteurs de* E *sera dite* linéairement dépendante *(ou* liée*) s'il existe des réels* $k_1, \ldots, k_m$ non tous nuls *tels que*

$$k_1.u_1 + \ldots + k_m.u_m = 0.$$

Exemples

1. Toute famille constituée d'un seul vecteur, autre que le vecteur nul, forme une famille libre.
2. Toute famille contenant le vecteur nulle est liée.
3. Une famille de deux vecteurs est liée si les deux vecteurs sont colinéaires (i.e. ils appartiennent à une même droite).
4. Dans le plan muni d'un point d'origine O, les vecteurs $\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}$ sont liés. Sur le dessin, une combinaison linéaire de ces vecteur se représente par une mise bout à bout de vecteurs colinéaires au vecteurs donnés. Or on voit qu'il en existe une qui revient au point O.



5. Dans $\mathbb{R}^3$, la famille $\{\overrightarrow{e_1} = (1,0,0), \overrightarrow{e_2} = (0,1,0), \overrightarrow{e_3} = (0,0,1), \overrightarrow{e_4} = (1,2,3)\}$ est liée, mais la sous-famille $\{\overrightarrow{e_1} = (1,0,0), \overrightarrow{e_2} = (0,1,0), \overrightarrow{e_3} = (0,0,1)\}$ est libre.
   En effet, remarquons tout d'abord que

$$\overrightarrow{e_4} = \overrightarrow{e_1} + 2.\overrightarrow{e_2} + 3.\overrightarrow{e_3}.$$

   La famille étudiée est donc bien liée.
   Montrons maintenant que la sous famille $\{\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3}\}$ est libre. Pour cela, considérons une combinaison linéaire de ces trois vecteurs qui s'annule. C'est à dire la donnée de trois réels α, β et γ tels que

$$\alpha.\overrightarrow{e_1} + \beta.\overrightarrow{e_2} + \gamma.\overrightarrow{e_3} = \overrightarrow{0}. \qquad (*)$$

   Pour montrer que notre famille est libre, on doit montrer que l'on a forcement $\alpha = \beta = \gamma = 0$. Or si l'on revient aux coordonnées dans $\mathbb{R}^3$, l'équation $(*)$ se lit

$$(\alpha, \beta, \gamma) = (0,0,0),$$

ce qui impose bien $\alpha = \beta = \gamma = 0$. □

Remarque : de façon générale, pour montrer qu'une famille de vecteurs $\{\overrightarrow{v_1}, \ldots, \overrightarrow{v_n}\}$ est libre, on montre que pour construire le vecteur nul comme une combinaison linéaire des $\overrightarrow{v_i}$ on doit prendre tous les coefficients nuls :

$$\alpha_1 \overrightarrow{v_1} + \ldots \alpha_n \overrightarrow{v_n} = \overrightarrow{0} \;\Rightarrow\; \alpha_i = 0 \, \forall i = 1..n$$

2.3.3 Bases et coordonnées

On appellera *b*ase d'un espace vectoriel E toute famille génératrice de E qui est optimale. Précisément,

Définition 2.3.4 *Étant donné un espace vectoriel E, une famille $\{\overrightarrow{v_1}, \ldots, \overrightarrow{v_n}\}$ de vecteurs de E est une* base de E *si c'est une famille libre qui engendre E.*

Grâce à tout ce que l'on vient de voir, il est clair qu'une base est ce que l'on appelais jusque là un repère. Ainsi, une fois que l'on connaît une base $\mathcal{B} = \{\overrightarrow{v_1}, \ldots, \overrightarrow{v_n}\}$ de notre espace vectoriel E, tout vecteur $\overrightarrow{u} \in E$ peut s'écrire de façon unique comme combinaison linéaire des $\overrightarrow{v_i}$

$$\overrightarrow{u} = k_1.\overrightarrow{v_1} + \ldots + k_n.\overrightarrow{v_n}.$$

Les coefficients $k_1, \ldots, k_n$ sont les coordonnées du vecteur $\overrightarrow{u}$ dans la base $\mathcal{B}$ et l'on note $\overrightarrow{u} = \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}_{\mathcal{B}}$.

Exemples

1. La famille $\mathcal{B} = \{(1,0),(1,1)\}$ est une base de $\mathbb{R}^2$. Ainsi, le vecteur $(3,2)$ s'écrit

   $$(3,2) = (1,0) + 2.(1,1).$$

   Les coordonnées de ce vecteur dans la base $\mathcal{B}$ sont donc

   $$(3,2) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}_{\mathcal{B}}.$$

   De façon générale, les coordonnées d'un vecteur (x,y) de $\mathbb{R}^2$ dans la base $\mathcal{B}$ sont

   $$(x,y) = \begin{pmatrix} x-y \\ y \end{pmatrix}_{\mathcal{B}}.$$

2. Dans $\mathbb{R}^3$, la famille $\{(1,1,0),(1,0,1),(0,1,1)\}$ est une base de $\mathbb{R}^3$.
   Cependant, la famille la sous-famille $\{(1,1,0),(1,0,1)\}$ n'est pas une base de $\mathbb{R}^3$, car elle n'est pas génératrice. Il manque un vecteur, une dimension. De même, la famille $\{(1,1,0),(1,0,1),(0,1,1),(1,2,3)\}$ n'est pas une base de $\mathbb{R}^3$ car elle n'est pas libre.

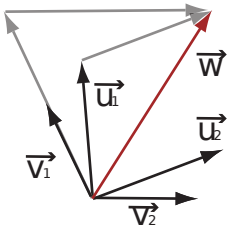
3. Soit n un entier fixé. Pour tout $i \in \{1, \dots, n\}$, on note $\overrightarrow{e_i}$ le vecteur de $\mathbb{R}^n$ dont toutes les coordonnées sont nulles, sauf la i-ième.

$$\overrightarrow{e_i} = (0, \dots, 0, \underset{\substack{\uparrow \\ i^{\text{ième}}}}{1}, 0, \dots, 0).$$

La famille $\mathcal{B}_n = \{\overrightarrow{e_1}, \dots, \overrightarrow{e_n}\}$ est alors une base de $\mathbb{R}^n$. C'est *la base canonique* de $\mathbb{R}^n$ car les coordonnées de n'importe quel vecteur $(x_1, \dots, x_n)$ de $\mathbb{R}^n$ dans la base $\mathcal{B}_n$ sont

$$(x_1, \dots, x_n) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}_n}.$$

4. Sur le dessin ci-contre, les famille $\mathcal{U} = \{\overrightarrow{u_1}, \overrightarrow{u_2}\}$ et $\mathcal{V} = \{\overrightarrow{v_1}, \overrightarrow{v_2}\}$ sont des bases du plan. Ainsi, le vecteur $\overrightarrow{w}$ s'écrit $\overrightarrow{w} = \overrightarrow{u_1} + \overrightarrow{u_2}$ et $w = 2.\overrightarrow{v_1} + 2\overrightarrow{v_2}$. On a donc



$$\overrightarrow{w} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\mathcal{U}} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}_{\mathcal{V}}.$$

2.3.4 Dimension d'un espace vectoriel

Définition

Pour terminer ce chapitre, revenons sur la notion de directions dans un espace vectoriel. L'une des caractéristiques importante que l'on peut observer concernant un espace vectoriel donné est le nombre de directions indépendantes que l'on peut trouver dans cet espace. Intuitivement, c'est cela que l'on appelle la dimension d'un EV (par exemple, $d = 2$ pour le plan, et $d = 3$ pour l'espace). Cependant, comme pour les autres notions associées à un EV, on souhaite en avoir une définition précise et rigoureuse. Elle nous est donnée via la notion de base.

Précisément, étant donné un espace vectoriel E, si l'on souhaite construire une famille génératrice de E, il faut que toutes les directions indépendantes de E soient représentées. Il faut donc que son cardinal soit au moins égal à la notion intuitive de dimension. (Si l'on se donne deux vecteurs dans l'espace, on ne pourra engendrer au maximum qu'un plan de l'espace).

D'autre part, si l'espace vectoriel E contient d directions *indépendantes*, cela signifie que si l'on se donne $d + 1$ vecteurs, il en est au moins un qui n'apporte pas de direction supplémentaire. Ce vecteur s'écrit donc comme combinaison linéaire des autres. Autrement dit, la famille n'est pas libre. On peut formaliser tout ça de la façon suivante.

Proposition 2.3.5 *Étant donné un espace vectoriel E, il existe un entier d tel que*

1. *toute famille génératrice de E possède* au moins *d vecteurs,*
2. *toute famille libre de E possède* au plus *d vecteurs,*
3. *toute base de E est composée de d vecteurs.*

C'est cet entier d que l'on appelle dimension *de E et que l'on note* $\dim E$.

Ainsi, en pratique, pour déterminer la dimension d'un espace vectoriel donné, on en cherche une base. Le nombre de vecteurs constituant cette base nous donne alors la dimension cherchée.

Réciproquement, si l'on se donne une famille $\mathcal{F}$ de n vecteurs dans E, la dimension du sous espace engendré par $\mathcal{F}$ correspond au nombre de directions indépendantes représentées. Autrement dit,

$$\dim(\mathrm{Vect}(\mathcal{F})) \leqslant n$$

(i.e. dans l'espace, deux vecteurs engendrent au plus un plan, dans $\mathbb{R}^4$, trois vecteurs engendrent au plus un sous espace de dimension 3).

Il est même possible d'être plus précis si la famille $\mathcal{F}$ est libre : si $\mathcal{F}$ est une famille libre, alors

$$\dim(\mathrm{Vect}(\mathcal{F})) = n$$

Cette dernière propriété réduits en particulier les calculs permettant de déterminer des bases dans un espace vectoriel dont on connaît la dimension.

Précisément, si E est un espace de dimension n, alors toutes famille libre de n vecteurs est nécessairement génératrice (les n directions indépendantes de E sont représentées). Ainsi, toute famille de deux vecteurs non colinéaire du plan est une base du plan, toute famille libre de trois vecteurs de l'espace engendre l'espace.

Notons que si la famille $\mathcal{F}$ est liée, on ne peut être plus précis que

$$\dim(\mathrm{Vect}(\mathcal{F})) < n$$

Pour déterminer cette dimension, il faut étudier plus en détails les dépendances linéaires qui lient la famille. On appelle alors *rang de* $\mathcal{F}$ la dimension de l'espace qu'elle engendre :

$$\mathrm{rg}(\mathcal{F}) \;=\; \dim(\mathrm{Vect}(\mathcal{F})).$$

Ainsi, dans l'espace, une famille de rang 2 engendre un plan. Tous les vecteurs de la famille sont coplanaires. Dans une famille de rang 1, tous les vecteurs sont sur la même ligne.

Pour finir, cette approche de la notion de vecteurs permet de démontrer les deux assertions suivantes :

Proposition 2.3.6 *Soit E un espace vectoriel et dimension n.*

1. *Toute famille libre peut être complétée en une base de E.*
2. *De toute famille génératrice de E, on peut extraire une base de E.*

2.4 Représentations matricielles

Les matrices offrent un outil très efficace dans tous les problèmes d'algèbre linéaire. Les opérations matricielles permettent en particulier de traduire toutes les notions que l'on vient de voir ainsi que tous les problèmes associés en termes d'équations matricielles. Une application importante de ces propriétés est le fait qu'une fois traduits en termes d'équations matricielles, les problèmes d'algèbre linéaire peuvent être confiés à une machine.

La première application que l'on va voir porte sur les familles de vecteurs. Les matrices permettent en effet de stocker et étudier les différents types de familles d'un même espace.

Avant d'en voir les applications, quelques rappels sur les matrices et le calcul matriciel.

2.4.1 Matrices

Définitions

Rappelons donc qu'une matrice est un tableau de nombres appelés coefficients. On note traditionnellement m son nombre de lignes et n son nombre de colonnes. Si $m = 1$, on parle de vecteur ligne ; si $n = 1$, on parle de vecteur colonne. Enfin, dans le cas $m = n$, on parle de matrice carrée.

Enfin, on repère chaque coefficient d'une matrice donnée par le numéro de sa ligne et celui de sa colonne : $A = (a_{ij})$.

Opérations matricielles

Addition. Si A et B sont deux matrices de même taille, on peut définir la somme $A + B$ en sommant coefficient par coefficient.

Multiplication. Si $A = (a_{ij})$ et $B = (b_{ij})$ sont deux matrices telles que le nombre n de colonnes de A est égal au nombre de lignes de B, on peut construire le produit $C = A \times B$ comme étant la matrice dont les coefficients c_{ij} sont définis par

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

On obtient ainsi une matrice dont le nombre de lignes est celui de A et dont le nombre de colonnes est celui de B.

Note : une particularité du produit matriciel est qu'il n'est pas commutatif. En effet, si A et B sont deux matrices telles que le produit $A \times B$ existe, ça n'est pas nécessairement le cas du produit $B \times A$. De plus, même quand les deux produits AB et BA existent, rien n'assure qu'ils soient égaux.

Invariants matriciels

À toute matrice $M \in \mathcal{M}_{m,n}(\mathbb{R})$, on peut associer certaines quantités caractéristiques. La plus emblématique est le déterminant sur lequel on reviendra au paragraphe suivant. Il s'agit d'un invariant scalaire (c'est un nombre).

Parmi les autres invariants scalaires, notons *le rang* d'une matrice qui, par analogie avec le rang d'une famille de vecteur est défini comme étant la dimension du sous espace de $\mathbb{R}^n$ engendré par la famille des lignes.

D'autre part, on peut associé à une matrice M donnée une autre matrice, appelée *transposée de* M et notée ${}^tM$. Par définition, ${}^tM$ est la matrice obtenue en inversant les lignes et les colonnes de M.

Matrices carrées

Une matrice M est dite carrée si elle a autant de lignes de que colonnes. Les matrices carrées vérifient certaines propriétés qui n'ont de sens que de ce cas. Notons en particulier que dans une matrice carrée, on peut distinguer les éléments dits *diagonaux*, i.e. les éléments dont le numéro de ligne et le numéro de colonne sont égaux : la diagonale de $A = (a_{ij})$ est constituée des termes a_{jj}.

On peut alors définir un nouvel invariant pour les matrices carrées : pour une matrice carrée A donnée, *la trace* de A, notée $\mathrm{Tr}(A)$ est par définition la somme de ses termes diagonaux.

D'autre part, on peut distinguer certaines sous familles dans la famille des matrices carrées.

Matrices symétriques. Si M est une matrice carrée, la transposée ${}^tM$ de M est obtenue par symétrie par rapport à sa diagonale. M est alors dite symétrique si cette symétrie ne change pas la matrice. Ainsi, M est symétrique si et seulement si ${}^tM = M$.

Matrices diagonales. On appelle matrice carrée toute matrice ayant autant de lignes que de colonnes. Dans une matrice carrée, on distingue en particulier *la diagonale*. Elle correspond aux éléments dont le numéro de ligne est égal au numéro de colonne. On appelle alors matrice diagonale toute matrice carrée dont les termes non diagonaux sont nuls. On

verra en fin de chapitre que les matrices diagonale jouent un rôle fondamental dans tout problème d'algèbre linéaire.

Matrices inverses. Parmi toutes les matrices diagonales, on appelle *matrice identité* la matrice diagonale dont tous les coefficients diagonaux valent 1 et on la note I_n (n étant le nombre de lignes ou de colonnes). Elle joue un rôle fondamental dans le cadre des opérations matricielles puisque pour toute matrice carrée A de taille n, on a

$$I_n \times A = A \times I_n = A.$$

Étant donnée une matrice carrée A, on appelle alors inverse de A la matrice notée A^{-1} telle que

$$A \times A^{-1} = A^{-1} \times A = I_n.$$

La notion de matrice inverse permet de résoudre des systèmes linéaires carrés. Cependant, toutes les matrices carrées ne sont par inversibles. Ainsi, pour qu'une matrice soit inversible, il faut (et il suffit) que son déterminant soit *non nul*.

Déterminant

Le déterminant d'une matrice est un outils permettant de connaître de nombreuse propriétés sur cette matrices ; mais dont le calcul n'est jamais simple (y compris à l'aide d'une machine).

Intuitivement, le déterminant permet (entre autres) de mesurer la dépendance linéaire d'une famille de vecteurs : la famille des vecteurs lignes (ou celle des vecteurs colonnes) et donc toutes les familles d'objets que l'on peut représenter à l'aide de coordonnées.

Ainsi, en deux variables, les deux lignes $u = \begin{pmatrix} a & b \end{pmatrix}$ et $v = \begin{pmatrix} c & d \end{pmatrix}$ de la matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ sont dépendantes (et forment donc une famille liée) si elles sont proportionelles. i.e. si les quotients $\frac{a}{c}$ et $\frac{b}{d}$ sont égaux. Or

$$\frac{a}{c} = \frac{b}{d} \Leftrightarrow ad = bc \Leftrightarrow ad - bc = 0.$$

On appelle alors déterminant de A la quantité

$$\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

2.4.2 Changement de base

Comme on vient de le voir dans l'exemple précédent, un espace vectoriel possède plusieurs bases (il en possède même une infinité...). Et les coordonnées d'un vecteur donné

changent si l'on change de base. D'un point de vue pratique, on est souvent amenés à changer de base au cours de l'étude d'un problème (on cherchera notamment à trouver des bases plus adaptées au problème, i.e. des bases dans lesquelles nos vecteurs ont des coordonnées plus simples...). Or le problème de changement de base se réduit souvent à un problème de multiplication de matrices.

Exemples de changement de base. On se place dans l'espace vectoriel $\mathbb{R}^3$, muni de sa base canonique

$$\mathcal{B}_1 = \{\overrightarrow{e_1} = (1,0,0), \overrightarrow{e_2} = (0,1,0), \overrightarrow{e_3} = (0,0,1)\}$$

et l'on considère le vecteur $\overrightarrow{v} = (3,1,-4)$. Ses coordonnées dans la base canonique sont $\begin{pmatrix} 3 \\ 1 \\ -4 \end{pmatrix}_{\mathcal{B}_1}$. Soit maintenant la famille $\mathcal{B}_2$ formée des vecteurs suivants.

$$\overrightarrow{f_1} = (1,1,1), \quad \overrightarrow{f_2} = (0,1,1) \text{ et } \overrightarrow{f_3} = (0,0,1).$$

On peut montrer que $\mathcal{B}_2$ est une base de $\mathbb{R}^3$ et l'on veut obtenir les coordonnées de $\overrightarrow{v}$ dans cette base.

On cherche donc $\begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\mathcal{B}_2}$ tel que $\overrightarrow{v} = x.\overrightarrow{f_1} + y.\overrightarrow{f_2} + z.\overrightarrow{f_3}$. D'où

$$\begin{aligned}(3,1,-4) &= x.(1,1,1) + y.(0,1,1) + z.(0,0,1)\\ &= (x,x,x) + (0,y,y) + (0,0,z)\\ &= (x, x+y, x+y+z).\end{aligned}$$

On en déduit le système de trois équations linéaires suivant vérifiées par x, y et z,

$$\begin{cases} x & = & 3 \\ x+y & = & 1 \\ x+y+z & = & -4 \end{cases}$$

et l'on en déduit les coordonnées recherchées. Remarquons que l'on peut donc réduire le problème à une équation matricielle :

$$\begin{pmatrix} 3 \\ 1 \\ -4 \end{pmatrix} = \begin{array}{c} \overrightarrow{f_1}\ \overrightarrow{f_2}\ \overrightarrow{f_3} \\ \downarrow\ \downarrow\ \downarrow \\ \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \end{array} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

anciennes coordonnées ↑ nouvelles coordonnées ↑

On peut montrer que cette formule est générale. Précisément,

Théorème 2.4.1 *Soient $\mathcal{B}_1$ et $\mathcal{B}_2$ deux bases de E et soit $\overrightarrow{v}$ un vecteur de E.*

On note $\overrightarrow{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}_1}$ les coordonnées de $\overrightarrow{v}$ dans $\mathcal{B}_1$ et l'on note $P_{\mathcal{B}_1,\mathcal{B}_2}$ la matrice dont les colonnes sont formées des cordonnées des vecteurs de $\mathcal{B}_2$ dans la base $\mathcal{B}_1$.

Alors les coordonnées $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{\mathcal{B}_2}$ de $\overrightarrow{v}$ dans $\mathcal{B}_2$ vérifient

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}_1} = P_{\mathcal{B}_1,\mathcal{B}_2} \times \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{\mathcal{B}_2}.$$

La matrice $P_{\mathcal{B}_1,\mathcal{B}_1}$ est appelé matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$.

2.5 Applications linéaires

2.5.1 Définition

Grâce au formalisme des espaces vectoriels, on sait notamment représenter les points ou des figures du plan ou de l'espace par des éléments de $\mathbb{R}^2$ ou $\mathbb{R}^3$. Dans ce chapitre, on va mettre en place des outils permettant de faire bouger des points et ces figures. Pour cela, on utilise des fonctions d'un espace vectoriel dans lui même, ou plus généralement d'un espace vectoriel dans un autre.

Ces fonctions nous permettent de déplacer nos figures point par point. Si l'on veut que la figure ne sera pas trop déformée pendent le voyage, on doit faire quelques restrictions sur les fonctions que l'on considère. On ne considère ainsi que les fonctions dites linéaires. La seule chose que l'on demandera à une fonction linéaire est qu'elle conserve les combinaisons linéaires. D'où :

Définition 2.5.1 *Soient F et E deux espaces vectoriels, et soit f une fonction $F \to V$. L'application f est dite* linéaire *si*

$$\begin{aligned} \forall \overrightarrow{u}, \overrightarrow{v} \in F, \quad & f(\overrightarrow{u} + \overrightarrow{v}) = f(\overrightarrow{u}) + f(\overrightarrow{v}), \\ \forall \overrightarrow{u} \in F, \forall k \in \mathbb{R}, \quad & f(k.\overrightarrow{u}) = k.f(\overrightarrow{u}). \end{aligned}$$

Autrement dit, une application linéaire est une application entre espaces vectoriels qui respecte la combinaison linéaire :

$$\forall \overrightarrow{u}, \overrightarrow{v} \in F, \forall k, \ell \in \mathbb{R}, \quad f(k.\overrightarrow{u} + \ell.\overrightarrow{v}) = k.f(\overrightarrow{u}) + \ell.f(\overrightarrow{v})$$

Exemples.

1. Toute application du type
$$\begin{array}{rccc} f_a & : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & ax \end{array}$$
est une application linéaire.
2. Les applications
$$\begin{array}{rccc} \varphi & : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & 2x+1 \end{array}, \qquad \begin{array}{rccc} \psi & : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & x^2 \end{array}$$
ne sont pas linéaires.
3. Toute application du type
$$\begin{array}{rccc} g & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & & (x,y) & \longmapsto & (ax+by, cx+dy) \end{array}$$
est une application linéaire.
4. La projection
$$\begin{array}{rccc} p & : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & & (x,y,z) & \longmapsto & (x,y,0) \end{array}$$
est linéaire.
5. Une rotation, dans le plan ou dans l'espace est une application linéaire.

2.5.2 Premières propriétés

Ainsi, en conservant les combinaisons linéaires, les applications linéaires respectent la structure d'espace vectoriel. Une première conséquence emblématique de cet propriété est que toute application linéaire envoie le vecteur nul sur le vecteur nul :
$$f(\overrightarrow{0}) = f(0.\overrightarrow{v}) = 0.f(\overrightarrow{v}) = \overrightarrow{0}.$$
(cela semble naturel étant donné le rôle du vecteur nul en algèbre linéaire).

De façon plus générale, une droite est (presque toujours) envoyée sur une droite. En effet, si D est une droite dirigée par un vecteur $\overrightarrow{v}$, les images des vecteurs $k.\overrightarrow{v}$ de la droite sont de la forme
$$f(k.\overrightarrow{v}) = k.f(\overrightarrow{v})$$
Ils sont donc tous sur la droite portée par $f(\overrightarrow{v})$. L'image d'une droite est donc une droite, à condition que $f(\overrightarrow{v})$ soit différent de $\overrightarrow{0}$ (dans ce dernier cas, tous les vecteurs de la droite portée par $\overrightarrow{v}$ ont pour image $\overrightarrow{0}$).

On peut montrer de même qu'un plan toujours envoyé soit sur un plan, soit sur une droite, soit sur le vecteur nul.

Nous reviendrons plus en détails sur l'étude de ces sous ensembles appelés *images* par f de sous espaces de E.

2.5.3 Matrice d'une application linéaire

Un autre caractère important des applications linéaires est que pour calculer l'image de n'importe quel vecteur $\overrightarrow{v}$, il suffit de connaître l'image d'une base $\mathcal{B}$ et les coordonnées de $\overrightarrow{v}$ dans cette base. En effet : La première chose que l'on peut noter est que pour connaître l'image de n'importe quel vecteur de l'espace vectoriel dans lequel on travaille par une application linéaire, il nous suffit de connaître les images des vecteurs d'une base. En effet, soit E un espace vectoriel et soit $\mathcal{B} = \{\overrightarrow{e_1}, \ldots, \overrightarrow{e_n}\}$ une base de E. Tout vecteur $\overrightarrow{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}} \in E$ s'écrit

$$\overrightarrow{v} = x_1.\overrightarrow{e_1} + \ldots + x_n.\overrightarrow{e_n}.$$

Si f est une application linéaire $E \to E$, on a donc

$$\begin{aligned} f(\overrightarrow{v}) &= f(x_1.\overrightarrow{e_1} + \ldots + x_n.\overrightarrow{e_n}) \\ &= x_1.f(\overrightarrow{e_1}) + \ldots .x_n.f(\overrightarrow{e_n}). \end{aligned}$$

Toute l'information portant sur la fonction f se trouve donc dans les images $f(\overrightarrow{e_i})$ des vecteurs de la base $\mathcal{B}$. Or puisque ces images $f(\overrightarrow{e_i})$ sont des vecteurs de E, ils ont des coordonnées dans la base $\mathcal{B}$:

$$\forall j = 1..n, \quad f(\overrightarrow{e_j}) = a_{1j}\overrightarrow{e_1} + a_{2j}\overrightarrow{e_2} + \ldots + a_{nj}\overrightarrow{e_n} = \sum_{i=1}^{n} a_{ij}\overrightarrow{e_i}, \qquad a_{ij} \in \mathbb{R}.$$

En rassemblant ces coordonnées dans une matrice

$$A = \begin{array}{c} \begin{array}{cccc} f(\overrightarrow{e_1}) & f(\overrightarrow{e_2}) & \cdots & f(\overrightarrow{e_n}) \\ \downarrow & \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ a_{n1} & \cdots & a_{n,n-1} & a_{nn} \end{pmatrix} \end{array}$$

on construit *la matrice de f dans la base $\mathcal{B}$*. On peut alors déterminer l'image de n'importe quelle vecteur dont on connaît les coordonnées dans $\mathcal{B}$:

$$\forall \overrightarrow{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}} \in E, \qquad f(\overrightarrow{v}) = A \times \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Note : le produit $A \times \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ est un vecteur. Il contient les coordonnées de $f(\overrightarrow{v})$ dans la base $\mathcal{B}$.

Enfin, la forme matricielle permet de stocker une fonction linéaire dans un ordinateur et de lui confier les calculs d'images.

2.5.4 Image et noyau d'une application linéaire

Image

Comme on l'a évoqué plus haut, il est souvent utile de connaître l'image d'un sous ensemble par une application linéaire donnée. On va voir que la théorie donne de nombreux outils pour déterminer ces images à partir de l'image d'une base.

Ainsi, soient donc E un espace vectoriel de dimension n, $\mathcal{B} = \{\overrightarrow{e_1}, \ldots, \overrightarrow{e_n}\}$ une base de E et $f : E \to E$ une application linéaire. On appelle *image* de l'application f l'ensemble des vecteurs de E qui sont l'image d'un vecteur de E par f. On note

$$\mathrm{Im}(f) = \{\overrightarrow{v} \in E \,/\, \exists \overrightarrow{u} \in E \text{ tq } \overrightarrow{v} = f(\overrightarrow{u})\}.$$

Le fait (évoqué plus haut) que l'image d'une droite soit une droite ou un point n'est qu'un cas particulier de la propriété suivante : l'image d'un espace vectoriel par une application linaire est un espace vectoriel.

En effet, montrons que si $f : E \to E$ est une application linéaire, alors $\mathrm{Im}(f)$ est un sous espace vectoriel de E :

- D'abord, puisque f est linéaire, on a nécessairement $f(\overrightarrow{0}) = \overrightarrow{0}$. Le vecteur $\overrightarrow{0}$ est donc l'image de $\overrightarrow{0} \in E$. Il est donc dans $\mathrm{Im}(f)$.

- D'autre part, soient $\overrightarrow{v_1}, \overrightarrow{v_2} \in \mathrm{Im}(f)$ et soient $k, \ell \in \mathbb{R}$. Puisque $\overrightarrow{v_1}$ et $\overrightarrow{v_2}$ sont dans l'image de f, il existe des vecteurs $\overrightarrow{u_1}$ et $\overrightarrow{u_2}$ tels que $f(\overrightarrow{u_1}) = \overrightarrow{v_1}$ et $f(\overrightarrow{u_2}) = \overrightarrow{v_2}$. Mais alors,

$$\begin{aligned} k.\overrightarrow{v_1} + \ell.\overrightarrow{v_2} &= k.f(\overrightarrow{u_1}) + \ell.f(\overrightarrow{u_2}) \\ &= f(k.\overrightarrow{u_1}) + f(\ell.\overrightarrow{u_2}) \\ &= f(k.\overrightarrow{u_1} + \ell.\overrightarrow{u_2}). \end{aligned}$$

  Le vecteur $k.\overrightarrow{v_1} + \ell.\overrightarrow{v_2}$ est donc l'image par f de $k.\overrightarrow{u_1} + \ell.\overrightarrow{u_2} \in E$ par f. C'est donc bien un vecteur de $\mathrm{Im}(f)$ et $\mathrm{Im}(f)$ est stable par combinaisons linéaires.

Puisque $\mathrm{Im}(f)$ est stable par combinaisons linéaires et qu'il contient le vecteur nul, c'est bien un SEV de E.

Or si l'on connaît une base $\mathcal{B} = \{\overrightarrow{e_1}, \ldots, \overrightarrow{e_n}\}$ de E, l'image par f des vecteurs de $\mathcal{B}$ (et donc la matrice de f dans $\mathcal{B}$) permet d'obtenir une base de $\mathrm{Im}(f)$ (et d'en connaître ainsi sa dimension). En effet, tout vecteur de $\mathrm{Im}(f)$ est une combinaison des vecteurs $f(\overrightarrow{e}_i)$: pour tout $\overrightarrow{v} \in \mathrm{Im}(f)$, il existe $\overrightarrow{u} \in E$ tel que $\overrightarrow{v} = f(\overrightarrow{u})$. Or comme $\overrightarrow{u} \in E$, $\overrightarrow{u}$ s'écrit comme combinaison linéaire des $\overrightarrow{e_i}$:

$$\overrightarrow{u} = x_1\overrightarrow{e_1} + \ldots + x_n\overrightarrow{e_n}.$$

Mais alors, $\overrightarrow{v} = f(\overrightarrow{u})$ s'écrit

$$\overrightarrow{v} = x_1 f(\overrightarrow{e_1}) + \ldots + x_n f(\overrightarrow{e_n}).$$

Le sous espace $\mathrm{Im}(f)$ est donc engendré par la famille $\mathcal{F} = \{f(\overrightarrow{e}_1), \ldots, f(\overrightarrow{e_n})\}$. Par analogie avec les familles de vecteurs, on appelle *rang de* f la dimension de $\mathrm{Im}(f)$ et puisque $\mathcal{F}$ contient n vecteurs, on a

$$\mathrm{rg}(f) \leqslant n.$$

Si la famille $\mathcal{F}$ est libre (i.e. si le rang de f est maximal), elle forme une base de $\mathrm{Im}(f)$. Dans ce cas, $\mathrm{Im}(f)$ est de dimension n et correspond à l'espace E tout entier. On dit alors que la fonction f est *surjective* (tous les vecteur de E sont des images).

Si la famille $\mathcal{F}$ n'est pas libre, on a $\mathrm{rg}(f) < n$. $\mathrm{Im}(f)$ est donc strictement plus petit que E et l'étude de la matrice de f dans $\mathcal{B}$ permet d'en déterminer la dimension ainsi qu'une base.

Noyau

Une autre propriété fondamentale des applications linéaires porte sur l'ensemble des vecteurs dont l'image est nulle. Ainsi, on appelle *noyau* de f l'ensemble $\mathrm{Ker}(f)$ des vecteurs de E qui sont envoyés sur $\overrightarrow{0}$:

$$\mathrm{Ker}(f) = \{\overrightarrow{u} \in E \,/\, f(\overrightarrow{u}) = \overrightarrow{0}\}.$$

Comme pour l'image, on vérifie dans un premier temps de le noyau d'une application est un espace vectoriel (i.e. un sous espace vectoriel de E) :

- Comme $f(\overrightarrow{0}) = \overrightarrow{0}$, le vecteur $\overrightarrow{0}$ est bien un élément du noyau de f.

- Soient $\overrightarrow{u_1}$ et $\overrightarrow{u_2}$ deux vecteurs de $\mathrm{Ker}(f)$ et soient $k, \ell \in \mathbb{R}$. Pour montrer que la combinaison $k.\overrightarrow{u_1} + \ell.\overrightarrow{u_2}$ est encore dans $\mathrm{Ker}(f)$, on vérifie que son image par f est $\overrightarrow{0}$:

$$\begin{aligned} f(k.\overrightarrow{u_1} + \ell.\overrightarrow{u_2}) &= f(k.\overrightarrow{u_1}) + f(\ell.\overrightarrow{u_2}) \\ &= k.f(\overrightarrow{u_1}) + \ell.f(\overrightarrow{u_2}) \\ &= k.\overrightarrow{0} + \ell.\overrightarrow{0} \;=\; \overrightarrow{0}. \end{aligned}$$

  $\mathrm{Ker}(f)$ est donc également stable par combinaisons linéaires.

En pratique, pour déterminer le noyau de f, on résout le système linéaire associé à l'équation $f(\overrightarrow{u}) = \overrightarrow{0}$. La taille du noyau $\mathrm{Ker}(f)$ est alors inversement proportionnel au nombre de contraintes présentes dans le système associé.

Théorème du rang

Soit $f : E \to E$ une application linéaire. On appelle *rang de* f la dimension du SEV $\mathrm{Im}(f)$ (que l'on note $\mathrm{rg}(f)$). On verra plus loin comment le rang de f, ainsi que la dimension de $\mathrm{Ker}(f)$ nous donnent des renseignements importants sur les propriétés de f. Il nous faut donc des outils permettant de calculer rapidement les dimensions de ces sous

espaces. Le théorème du rang est une égalité particulièrement efficace reliant les dimensions de $\mathrm{Im}(f)$ et $\mathrm{Ker}(f)$, ce qui nous permettra de déterminer ces dimensions plus rapidement qu'en produisant une base. Mais avant de l'énoncer, voyons quelques inégalités "naïves" mais néanmoins utiles :

- Tout d'abord, comme $\mathrm{Ker}(f)$ est un SEV de E, on a clairement
$$\dim(\mathrm{Ker}(f)) \leqslant \dim E.$$
- De même, comme $\mathrm{Im}(f) \subset E$, on a
$$\mathrm{rg}(f) \leqslant \dim E.$$
- Enfin, comme $\mathrm{Im}(f)$ est engendrée par l'image des vecteurs d'une base de E, on a
$$\mathrm{rg}(f) \leqslant \dim E.$$

Mais le résultat le plus fort concernant ces dimensions est le suivant :

Théorème 2.5.2 *Soit f une application linéaire de E dans E. Alors*
$$rg(f) + \dim(Ker(f)) = \dim E.$$

Ce résultat nous permettra souvent de n'étudier que le noyau ou que l'image et d'en tirer des informations sur les deux sous espaces et sur la fonction elle même.

Preuve : notons $n = \dim E$ et $p = \dim(\mathrm{Ker}(f))$. Comme $\mathrm{Ker}(f)$ est un SEV de E, on a $p \leqslant n$ et il existe une famille libre $\{\overrightarrow{e_1}, \ldots, \overrightarrow{e_p}\}$ de vecteur de E qui forme une base de $\mathrm{Ker}(f)$. On peut alors compléter cette famille en une base $\mathcal{B} = \{\overrightarrow{e_1}, \ldots, \overrightarrow{e_p}, \overrightarrow{e_{p+1}}, \ldots, \overrightarrow{e_n}\}$ en ajoutant $n - p$ vecteur bien choisis.

Mais alors, pour tout vecteur $\overrightarrow{v} \in \mathrm{Im}(f)$, il existe $\overrightarrow{u} \in E$ tel que $\overrightarrow{v} = f(\overrightarrow{u})$. Or si $\overrightarrow{u} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}$, on a

$$\begin{aligned} \overrightarrow{v} &= f(x_1\overrightarrow{e_1} + \ldots + x_p\overrightarrow{e_p}, \ldots, x_n\overrightarrow{e_n}) \\ &= x_1 f(\overrightarrow{e_1} + \ldots + x_p f(\overrightarrow{e_p}) + x_{p+1} f(\overrightarrow{e_{p+1}}) + \ldots + x_n f(\overrightarrow{e_n}) \\ &= x_{p+1} f(\overrightarrow{e_{p+1}}) + \ldots + x_n f(\overrightarrow{e_n}) \end{aligned}$$

car les p premiers vecteurs sont envoyés sur 0. On peut en conclure que $\mathrm{Im}(f)$ est engendrée par la famille $\mathcal{B}' = \{f(\overrightarrow{e_{p+1}}), \ldots, f(\overrightarrow{e_n})\}$. Si l'on parvient à montrer que cette famille est libre, on aura une base de $\mathrm{Im}(f)$. On aura alors

$$\begin{aligned} \mathrm{rg}(f) &= \dim(\mathrm{Im}(f)) = \mathrm{Card}(\mathcal{B}') \\ &= n - p = \dim E - \dim(\mathrm{Ker}(f)). \end{aligned}$$

Montrons donc que la famille $\mathcal{B}'$ est libre : soit $\lambda_{p+1}, \ldots, \lambda_n$ des coefficients tels que

$$\lambda_{p+1} f(e_{p+1}) + \ldots + \lambda_n f(\overrightarrow{e_n}) = \overrightarrow{0}_E.$$

Comme f est linéaire, on a donc

$$f(\lambda_{p+1} e_{p+1} + \ldots + \lambda_n \overrightarrow{e_n}) = \overrightarrow{0}_E.$$

Autrement dit, le vecteur $\lambda_{p+1} e_{p+1} + \ldots + \lambda_n \overrightarrow{e_n}$ est dans le noyau de f. Ce vecteur a donc des coordonnées $\lambda_1, \ldots, \lambda_p$ dans $\mathcal{B}$:

$$\begin{aligned} &\lambda_{p+1} e_{p+1} + \ldots + \lambda_n \overrightarrow{e_n} = \lambda_1 e_1 + \ldots + \lambda_p \overrightarrow{e_p} \\ \Leftrightarrow &\lambda_1 e_1 + \ldots + \lambda_p \overrightarrow{e_p} - \lambda_{p+1} e_{p+1} - \ldots - \lambda_n \overrightarrow{e_n} = \overrightarrow{0}. \end{aligned}$$

On a donc une combinaison linéaire des $\overrightarrow{e_i}$ qui revient au vecteur nul. Comme cette famille est libre (c'est une base de F) on doit avoir $\lambda_i = 0$ pour tout i, et en particulier les coefficients $\lambda_{p+1} = \ldots = \lambda_n = 0$. Autrement dit, la famille $f(\overrightarrow{e_{p+1}}), \ldots, f(\overrightarrow{e_n})$ est libre. C'est donc une base de $\mathrm{Im}(f)$ et le théorème du rang est démontré.

2.5.5 Injectivité, surjectivité d'une application linéaire

Rappels

Soient F et E deux ensembles.

1. Une fonction $f : F \to E$ est dite *injective* si deux éléments distincts de F sont toujours envoyés sur deux éléments distincts de E (i.e. si "toutes les images sont différentes") :

   $$(u \neq u') \Rightarrow (f(u) \neq f(u'))$$

   ou

   $$(f(u) = f(u')) \Rightarrow (u = u').$$

2. Une fonction $f : F \to E$ est dite *surjective* si chaque élément de E peut s'écrire comme l'image d'un élément de F (i.e. si tous les éléments à l'arrivée sont atteints) :

   $$\forall v \in E, \exists u \in F \text{ tq } v = f(u).$$

3. Une fonction f qui est à la fois injective et surjective sera dite *bijective*. On peut alors définir sa fonction inverse $f^{-1} : E \to F$ définie par

   $$\begin{aligned} &\forall u \in F, \quad f^{-1} \circ f(u) = u, \\ &\forall v \in E, \quad f \circ f^{-1}(v) = v \end{aligned}$$

Injectivité, surjectivité des applications linéaires

Si l'on s'intéresse à l'injectivité ou la surjectivité d'une application *linéaire* $f : F \to E$, l'étude de l'image et du noyau nous permettent rapidement de régler le problème.

– Il est tout d'abord clair que l'étude de $\mathrm{Im}(f)$ nous renseigne sur la surjectivité de f. En effet, savoir si l'ensemble des vecteurs de E sont atteints revient à savoir si $\mathrm{Im}(f)$ recouvre E tout entier ou non. Or comme on sait déjà que $\mathrm{Im}(f)$ est un SEV de E, il nous suffit de déterminer la dimension de $\mathrm{Im}(f)$. Si elle correspond à la dimension de E, alors f est surjective. Sinon (i.e. si elle est strictement plus petite), f n'est pas surjective.

  Remarque : pour toute fonction linéaire f, le SEV $\mathrm{Im}(f)$ est engendré par les images de n'importe quelle base de F. Ainsi, si $m = \dim(F)$, $\mathrm{Im}(f)$ est engendré par m vecteurs. Il est donc de dimension au plus m. Ainsi, si E est trop grand, i.e. si $\dim(E) > \dim(F)$, on ne pourra pas construire d'application linéaire $F \to E$ surjective.

– Voyons maintenant comment l'étude du noyau de f nous renseigne sur l'injectivité de f. Comme on vient de le rappeler, une fonction est injective si

$$\forall \overrightarrow{u}, \overrightarrow{v} \in E, \qquad (\, f(\overrightarrow{u}) = f(\overrightarrow{v}) \,) \Leftrightarrow (\, \overrightarrow{u} = \overrightarrow{v} \,).$$

  Or si f est linéaire, cette condition s'écrit également

$$\forall \overrightarrow{u}, \overrightarrow{v} \in E, \qquad (\, f(\overrightarrow{u} - \overrightarrow{v}) = \overrightarrow{0} \,) \Leftrightarrow (\, \overrightarrow{u} - \overrightarrow{v} = \overrightarrow{0} \,).$$

  ou

$$\forall \overrightarrow{v} \in E, \qquad (\, f(\overrightarrow{v}) = \overrightarrow{0} \,) \Leftrightarrow (\, \overrightarrow{v} = \overrightarrow{0} \,).$$

  Autrement dit, une application linéaire est injective si son noyau ne contient que le vecteur nul. Là encore, il nous suffit donc de déterminer la dimension de $\mathrm{Ker}(f)$ pour déterminer si f est injective ou non.

Ainsi, déterminer si une fonction linéaire f est injective ou surjective (et donc bijective) n'est donc qu'une histoire de dimension. On verra plus loin des outils (de calcul matriciel) nous permettant de répondre rapidement à la question.

On peut cependant déjà noter que le théorème du rang qui relie les dimensions du noyau et de l'image nous permet d'établir un résultat étonnant pour certaines applications linéaires. Précisément, si f est une application d'un EV E dans lui-même (ou plus généralement si F et E ont même dimension), alors soit f est bijective, soit elle est ni injective, ni surjective. Autrement dit, si f est injective ou surjective, alors elle est bijective. En effet, le théorème du rang nous dit que

$$\mathrm{rg}(f) + \dim(\mathrm{Ker}(f)) = \dim E.$$

Donc si f est injective, on a $\mathrm{Ker}(f) = \{\overrightarrow{0}\}$ et donc $\dim(\mathrm{Ker}(f)) = 0$. Mais alors $\mathrm{rg}(f) = \dim E$ et f est surjective, donc bijective.

Dans l'autre sens, si f est surjective, alors $\mathrm{Im}(f) = E$ et $\mathrm{rg}(f) = \dim E$. On a alors $\dim(\mathrm{Ker}(f)) = 0$. Donc $\mathrm{Ker}(f) = \{\overrightarrow{0}\}$ et f est injective, donc surjective.

On peut alors énoncer le théorème suivant :

Théorème 2.5.3 *Soit $f : E \to E$ une application linéaire. Alors*

$$f \text{ est injective } \Leftrightarrow f \text{ est surjective } \Leftrightarrow f \text{ est bijective.}$$

Exemple : soit $\begin{array}{rccc} f & : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & & (x,y,z) & \longmapsto & (x+2y-4z, y-z, 2z) \end{array}$

1. Vérifier rapidement que f est linéaire.
2. Déterminer le noyau de f.
3. En déduire que f est une bijection.
4. Déterminer l'antécédent de $(1,1,2)$, de $(0,3,-3)$.
5. Déterminer l'antécédent de n'importe quel vecteur $\overrightarrow{v} = (a,b,c)$.
6. En déduire une matrice de l'application réciproque f^{-1}.

2.5.6 Matrice d'une application linéaire

Définition

Revenons aux applications linéaires. Soit donc E un espace vectoriel et $f : E \longrightarrow E$ une application linéaire sur E. Si l'on connaît une base $\mathcal{B} = \{\overrightarrow{e_1}, \ldots, \overrightarrow{e_n}\}$ de E, on a vu que l'on pouvait rapidement calculer l'image par de n'importe quel vecteur $\overrightarrow{v} \in E$ (exprimé dans $\mathcal{B}$) une fois que l'on a calculer les images $f(\overrightarrow{e_i})$. Précisément, si

$$\overrightarrow{v} = x_1\overrightarrow{e_1} + \ldots + x_n\overrightarrow{e_n},$$

alors

$$f(\overrightarrow{v}) = x_1 f(\overrightarrow{e_1}) + \ldots + x_n f(\overrightarrow{e_n}). \qquad (*)$$

Par ailleurs, les vecteurs $f(\overrightarrow{e_i})$ sont également des vecteurs de E. Ils ont donc des coordonnées dans la base $\mathcal{B}$:

$$f(\overrightarrow{e_i}) = a_{1i}\overrightarrow{e_1} + \ldots a_{ni}\overrightarrow{e_n}.$$

En notant $(y_1, \ldots, y_n)$ les coordonnées de $f(\overrightarrow{v})$ dans la base $\mathcal{B}$, on peut écrire l'équation $(*)$ sous la forme d'un système :

$$\left\{ \begin{array}{rcl} y_1 & = & a_{11}x_1 + \ldots + a_{1n}x_n \\ y_2 & = & a_{21}x_1 + \ldots + a_{2n}x_n \\ \vdots & & \vdots \\ y_n & = & a_{n1}x_1 + \ldots + a_{nn}x_n \end{array} \right.$$

Autrement dit, en notant $\overrightarrow{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{B}}$ et $f(\overrightarrow{v}) = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{\mathcal{B}}$, on a

$$f(\overrightarrow{v}) = A \times \overrightarrow{v}$$

où A est la matrices des coordonnées des images $f(\overrightarrow{e_1}, \ldots, \overrightarrow{e_n})$ dans $\mathcal{B}$. On dit que A est la matrice de f dans la base $\mathcal{B}$ (notée aussi $M_{\mathcal{B}}(f)$) :

$$M_{\mathcal{B}}(f) = \begin{array}{c} \begin{array}{cccc} f(\overrightarrow{e_1}) & f(\overrightarrow{e_2}) & \cdots & f(\overrightarrow{e_n}) \\ \downarrow & \downarrow & & \downarrow \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \end{array} \begin{array}{c} \overrightarrow{e_1} \\ \overrightarrow{e_2} \\ \vdots \\ \overrightarrow{e_n} \end{array}$$

Autrement dit, une fois que l'on s'est fixé une base de E, toute application linéaire $E \to E$ peut s'écrire sous la forme d'une matrice.

Exemple : Plaçons nous dans $\mathbb{R}^2$ et notons $\mathcal{B} = \{\overrightarrow{e_1}, \overrightarrow{e_2}\}$ la base canonique de $\mathbb{R}^2$. (On a donc $\overrightarrow{e_1} = (1, 0)$ et $\overrightarrow{e_2} = (0, 1)$). Considérons alors la fonction

$$\begin{array}{rccc} f & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & & (x, y) & \longmapsto & (2x + y, x + 2y) \end{array}$$

1. Calculer $f(\overrightarrow{e_1})$ et $f(\overrightarrow{e_2})$.
2. Donner leurs coordonnées dans $\mathcal{B}$ (ici, c'est immédiat puisque $\mathcal{B}$ est la base canonique).
3. En déduire la matrice de f dans la base $\mathcal{B}$.

Remarques :

– On peut faire la même chose pour $f : F \to V$ en se fixant une base $\mathcal{B}_F$ de F et une base $\mathcal{B}_E$ de E.

– La matrice d'une application linéaire dépend totalement de la base que l'on s'est fixé. Autrement dit, changer de base revient à changer de matrice. Si l'on reprend l'exemple de l'exercice précédent, on peut rapidement montrer que $\mathcal{B}' = \{(1, 1), (0, 1)\}$ est une autre base de $\mathbb{R}^2$. On peut donc calculer la matrice de f dans cette base. On trouve

$$M_{\mathcal{B}'}(f) = \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}.$$

Notons qu'en notant $\overrightarrow{v_1} = (1, 1)$, la première colonne de cette nouvelle matrice se lit $f(\overrightarrow{v_1}) = 3\overrightarrow{v_1}$. Autrement dit, le vecteur $\overrightarrow{v_1}$ est colinéaire à son image. Cela rend la

colonne particulièrement simple, puis qu'elle n'a qu'un coefficient non nul.

– On appelle application identité de E (notés Id_E) l'application

$$\begin{array}{rccc} \mathrm{Id}_E & : & E & \longrightarrow & E \\ & & \overrightarrow{v} & \longmapsto & \overrightarrow{v} \end{array}.$$

On peut noter que quelque soit la base $\mathcal{B}$ que l'on se fixe, on a

$$M_{\mathcal{B}}(\mathrm{Id}_E) = I_n.$$

Composée de fonctions et produit de matrices

La première chose que l'on peut noter concernant la représentation matricielle des applications linéaire est qu'elle est compatible avec la composée d'applications linéaire. Précisément, fixons une base $\mathcal{B}$ de E et considérons deux fonctions linéaires f et g de E dans E. On peut construire l'application composée

$$\begin{array}{rccc} f \circ g & : & E & \longrightarrow & E \\ & & \overrightarrow{v} & \longmapsto & f(g(\overrightarrow{v})) \end{array}$$

C'est une application linéaire. (Á démontrer...). On peut alors montrer que

$$M_{\mathcal{B}}(f \circ g) \;=\; M_{\mathcal{B}}(f) \times M_{\mathcal{B}}(g).$$

(Exercice : s'en convaincre).

En particulier, on sait qu'une application f est bijective s'il existe une fonction f^{-1} telle que $f \circ f^{-1} = \mathrm{Id}$. D'un point de vue matriciel, f est bijective s'il existe une matrice M' telle que

$$M_{\mathcal{B}}(f) \times M' = I_n.$$

Autrement dit, f est une bijection si sa matrice $M_{\mathcal{B}}(f)$ dans $\mathcal{B}$ est inversible. On a alors

$$M_{\mathcal{B}}(f^{-1}) = (M_{\mathcal{B}}(f))^{-1}.$$

De même, la multiplication de matrices permet de représenter rapidement l'application successive d'une même fonction (une rotation, par exemple). Ainsi, si l'on note $f^n = f \circ f \circ \ldots \circ f$, on a

$$M_{\mathcal{B}}(f^n) = (M_{\mathcal{B}}(f))^n.$$

Le calcul des puissances d'une matrice est un outil de base de l'utilisation matricielle. Il n'y a pas de méthode simple pour calculer les puissances d'une vitesse. On verra plus loin l'une des méthodes les plus efficaces (notamment d'un point de vue algorithmique).

Exemple : on se place dans le plan muni d'un point d'origine $\mathcal{O}$ et deux vecteurs de base $\overrightarrow{i}$ et $\overrightarrow{j}$ orthogonaux. On note s l'application du plan dont la matrice dans la base $\mathcal{B} = \{\overrightarrow{i}, \overrightarrow{j}\}$ est

$$M_{\mathcal{B}}(s) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

1. Calculer $s(\overrightarrow{i})$ et $s(\overrightarrow{j})$.
2. Donner l'image de tout vecteur $\overrightarrow{v} = (x, y)_{\mathcal{B}}$.
3. Vérifier que $s \circ s = \mathrm{Id}$.
4. Identifier l'application s d'un point de vue géométrique.

Changement de base pour une application linéaire

Comme on l'a déjà dit, la représentation matricielle d'une application linéaire est intimement liée à la base que l'on s'est fixé au départ. Or on a vu que pour une fonction donnée, il pouvait y avoir une base plus adaptée que les autres pour représenter notre fonction. Si en particulier on arrive à trouver une base de E formée de vecteurs qui sont tous colinéaires à leur image par f, la matrice de f dans cette base st particulièrement simple. Précisément, si l'on note $\mathcal{B}' = \{\overrightarrow{\varepsilon_1}, \ldots, \overrightarrow{\varepsilon_n}\}$ une telle base, il existe des coefficients $\lambda_1, \ldots, \lambda_n$ tels que

$$\forall i = 1..n, \quad f(\overrightarrow{\varepsilon_i}) = \lambda_i \overrightarrow{\varepsilon_i}.$$

Autrement dit,

$$f(\overrightarrow{\varepsilon_1}) = \begin{pmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{\mathcal{B}'}, \quad f(\overrightarrow{\varepsilon_2}) = \begin{pmatrix} 0 \\ \lambda_2 \\ \vdots \\ 0 \end{pmatrix}_{\mathcal{B}'}, \quad \cdots, \quad f(\overrightarrow{\varepsilon_n}) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_n \end{pmatrix}_{\mathcal{B}'}.$$

La matrice de f dans cette base est donc

$$M_{\mathcal{B}'}(f) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

On verra plus loin comment trouver une telle base (quand elle existe).

Avant cela, voyons comment, à partir de la matrice de f dans une base donner, on peut calculer la matrice de f dans une autre base. Soient donc $\mathcal{B} = \{\overrightarrow{e_1}, \ldots, \overrightarrow{e_n}\}$ et $\mathcal{B}' = \{\overrightarrow{\varepsilon_1}, \ldots, \overrightarrow{\varepsilon_n}\}$ deux bases de E et soit P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.

Question : comment déterminer la matrice de $M_{\mathcal{B}'}(f)$ à partir de $M_{\mathcal{B}}(f)$?

Comme tout problème d'algèbre linéaire, on peut traduire la question en termes de systèmes linéaires. Ainsi, pour construire $M_{\mathcal{B}'}(f)$, il faut calculer les coordonnées des images $f(\overrightarrow{\varepsilon_i})$ dans $\mathcal{B}'$. Or

- la matrice $M_{\mathcal{B}}(f)$ permet de calculer les images $f(\overrightarrow{\varepsilon_i})$ en fonction des $\overrightarrow{e_j}$,
- la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ permet d'exprimer les $\overrightarrow{e_i}$ en fonction des $\overrightarrow{\varepsilon_i}$.

On peut alors montrer (toujours en transformant le problème en un système linéaire et en décortiquant les produit de matrices qui apparaissent) que

$$M_{\mathcal{B}'}(f) = P^{-1} \times M_{\mathcal{B}}(f) \times P. \qquad (\Leftrightarrow\ M' = P^{-1}MP).$$

2.5.7 Diagonalisation d'une application linéaire

Soit $f : E \to E$, $\mathcal{B} = \{\overrightarrow{e_1}, \ldots, \overrightarrow{e_n}\}$ une base quelconque de E et A la matrice de f dans la base $\mathcal{B}$. Diagonaliser f (ou A), c'est déterminer une base de E dans laquelle la matrice de f est diagonale. Pour cela, on a deux étapes :

1. Trouver la matrice diagonale D, c'est-à-dire trouver les valeurs $\lambda_1, \ldots, \lambda_n$ à mettre sur la diagonale.
2. Trouver n vecteurs $\overrightarrow{v_1}, \ldots, \overrightarrow{v_n}$ à mettre en colonnes pour former la matrice P.

En pratique, on fait ces étapes l'une après l'autre, dans cet ordre.

Valeurs propres

Les coefficients $\lambda_1, \ldots, \lambda_n$ sont appelés *valeurs propres* de la matrice A. Ce sont les valeurs de λ pour lesquelles

$$\det(A - \lambda.I_n) = 0.$$

Remarque : la matrice $A - \lambda.I_n$ est la matrice A à laquelle on soustrait λ sur la diagonale. Les valeurs propres de $\lambda_1, \ldots, \lambda_n$ sont dons les valeurs de λ qui rendent cette matrice non inversible.

En pratique, on calcule donc le déterminant de $A - \lambda.I_n$. On obtient un polynôme de degré n en λ appelé *polynôme caractéristique de la matrice* A. Il suffit alors de déterminer les racines de ce polynôme. Dans le cas $n = 2$, on sait faire. Dans les cas $n \geqslant 3$, c'est plus compliqué. Il faut en général être astucieux dans le calcul du déterminant et tenter de le factoriser au fur et à mesure du calcul. En particulier, la méthode de Sarrus est ici particulièrement déconseillée puisqu'elle donne le polynôme sous forme totalement développée.

Vecteurs propres

Les vecteurs propres sont précisément les vecteurs dont les coordonnées forment les colonnes de la matrice P. En pratique, chaque vecteur propre est associé à une valeur

propre. Précisément, les vecteurs propres associés à la valeur propre λ_i sont les solutions non nulles du système

$$A \times X = \lambda_i.X \quad (*).$$

Notons que cette équation matricielle s'écrit

$$(A - \lambda_i.I_n) \times X = 0.$$

Or les λ_i sont précisément les valeurs de λ pour lesquelles cette équation admet des solutions non nulles.

Pour une valeur de λ_i fixée, l'ensemble des solutions de l'équation $(*)$ est appelé sous espace propre associé à λ_i.

En pratique, une fois que l'on a déterminer les valeurs propres λ_i, on résout donc le système $(*)$ pour chacun des λ_i.

Pour construire la matrice P inversible, il nous faut alors trouver n vecteurs propres distincts qui forment une famille libre. Si c'est possible, on prend tous ces vecteurs, on les met en colonne et on obtient notre matrice P. On peut ensuite vérifier, en calculant P^{-1} que l'on a bien

$$P^{-1}AP = D,$$

D étant la matrice diagonale dont les coefficients diagonaux sont les λ_i.

Remarques :

1. On peut montrer que les sous espaces propres sont des SEV de l'espace $\mathbb{R}^n$ (i.e. des plans ou des droites).
2. On peut montrer que deux sous espaces propres différents ne se rencontrent qu'en 0.
3. Pour pouvoir diagonaliser D (i.e. pour pouvoir construire la matrice P inversible), il faut donc que les sous espaces propres de A soient "assez gros". Précisément, il faut qu'à eux tous, ils permettent de recouvrir l'espace $\mathbb{R}^n$ tout entier. Dans l'espace par exemple, il nous faut soit trois droite, soit un plan et une droite.

   Si ça n'est pas le cas, c'est que A n'est pas diagonalisable.

Exemple : diagonaliser la matrice

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$